GROWTH AND YIELD MODELLING OF *PINUS RADIATA* IN CANTERBURY, NEW ZEALAND

A thesis
submitted in partial fulfilment
of the requirement for the degree of
Doctor of Philosophy in Forestry
in the University of Canterbury

by

Weizhong Zhao

University of Canterbury

1999

UNIVERSITY OF CANTERBURY

CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

During this study two growth and yield models were developed for *Pinus radiata* D.Don plantations in Canterbury, New Zealand, namely CanSPBL and CanJuv. CanSPBL is a model for the whole rotation of stands owned by the Selwyn Plantation Board Limited in Canterbury. CantJuv is a model for juvenile growth from after planting to before thinning in Canterbury. An existing stand growth and yield model CANTY was examined using a newly established relational database of PSP measurements. Projection bias was shown for mean top height, basal area per hectare and volume per hectare.

Height-diameter relationships were modelled for estimating individual tree height and mean top height. Of the sixteen functional forms evaluated the Petterson equation with exponent –5 and the two-parameter Richards' equation led to the smallest mean square error at stand level. Incorporating stand age, site index and altitude into the selected Petterson equation reduced the mean square error by 72% for a pooled regional data set.

The new model, CanSPBL, was achieved with precision at both stand and tree level. The components of the stand model include mean top height, basal area per hectare, stems per hectare, volume per hectare and diameter distribution. The inclusion of altitude into a chosen polymorphic Schumacher difference equation significantly reduced mean square error by 17% for mean top height and 41% for basal area. An examination of this stand model using two sources of data at plot and stand levels showed little apparent bias.

Two main model components of tree diameter and tree mortality were developed to complete an individual-tree projection system. For projection of diameter, an approach based on relative basal area was found best, compared with many sigmoid difference equations. A logistic regression procedure showed that relative diameter, altitude, the interaction between initial stocking and projection interval length, and site index were significantly correlated to tree mortality.

Nine sets of experimental data and thirty-one temporary plots were used to refine the juvenile growth model of stand level, CanJuv. The components of the model were mean top height, basal area, sectional area at ground level, and survival. Regression analyses identified age, annual rainfall and weeding as the significant factors in determining growth and mortality.

CHAPTER 1

INTRODUCTION

1.1 BACKGROUND

This study covers a broad research area in forest modelling systems, that includes establishment and assessment of a database for Permanent Sample Plots (PSP), modelling of height-diameter relationships, evaluation of inventory sampling practice, validation of an existing stand growth and yield model CANTY, and development of growth and yield models from time of establishment to harvest age. More weight was put on developing two growth and yield models of *Pinus radiata* D.Don plantations: one named "CanSPBL", was aimed at projecting appropriate future yield within the whole rotation of plantations owned by the Selwyn Plantation Board Limited (SPBL); the other named "CanJuv", was aimed at calibrating response of juvenile trees to silvicultural operations and environmental differences in Canterbury.

The main data sets used in building model CanSPBL were PSP measurements in SPBL in Canterbury, New Zealand. They represent approximately 10 000 ha of stocked pure and even-aged radiata pine plantations owned by the company (SPBL, 1998). These plantations are scattered between the Waimakariri and Rakaia Rivers in Mid-Canterbury and they cover plains, hills and coastal sands. For juvenile growth modelling, however, the population data set covered North-Canterbury and Mid-Canterbury.

An existing stand growth model in Canterbury, named CANTY (Goulding, 1995), was suggested by SPBL and was shown in a preliminary screening in this study to be biased in projecting stand yields for SPBL's forests. The model, therefore, needed to be examined thoroughly to determine the source and magnitude of discrepancies between predictions and observations. A new model needed to be established to provide more precise and more detailed information. A computer simulation program for projecting future yield was required by SPBL.

When the study started, the majority of PSP records in SPBL were in a form of hard copy (paper copy) and a few were in soft copy of Quattro Pro files, so a database had to be established and common variables for stands and plots had to be computed. To estimate mean top height and individual-tree height or mean height for each diameter class appropriately, a study determining the best height-diameter equation forms was needed. The Petterson equation (Schmidt, 1967) is widely used in New Zealand to represent the height-diameter relationship (McEwen, 1978; Goulding, 1995) but its fitting performance for radiata pine needs to be examined.

A new model, which is more accurate and versatile for use from stands to individual trees, had to be developed. Managers need both stand models and individual-tree models to make decisions. Two levels of model should be regarded as complementary systems rather than mutually exclusive options (Burkhart, 1977; Leary, 1979; Daniels and Burkhart, 1988).

Juvenile growth, from planting to just before thinning for *Pinus radiata* D.Don, is usually ignored in existing models but managers are interested in knowing the juvenile growth in response to their early silvicultural investments (Mason, 1992). There was no juvenile growth model available for Canterbury except the work done by Zhao and Mason (1996). Modelling work was needed for research and for practical management.

1.2 Research Objectives

The following study objectives were set out based on the state of research on modelling systems, the requirements of the forestry company of SPBL, and the research work already done on juvenile growth modelling.

1) To establish and assess a PSP database for growth modelling.

It was necessary to establish a database for PSP measurements detailed at a tree level. Data Base Management Systems (DBMS) are better than text or spreadsheet based systems in data entry, data storage and query making. Variables required for growth modelling were to be computed through data manipulation procedures within the database. The main variables were to be screened to ensure that the quality of the database was appropriate for growth modelling.

2) To model the relationship between tree height and diameter.

Height-diameter equations, which pass through the origin (d=0 cm, h=1.40 m), attain an asymptote for large values of diameter and always have a positive slope, have been recognised as the most appropriate form. Linearity is another good property of equations, especially in permanent sample plot (PSP) database computational systems, because coefficients can be solved explicitly and uniquely. Robust equation forms which fit well under a wide range of stands are desirable.

At a stand level, various functional forms were to be evaluated for stands varying considerably in age, site index, altitude and the number of trees sampled. The best equation form, which is linear or could be transformed into linear, was needed to obtain regression coefficients and to estimate tree-height of each stand measurement within a database system. A prediction model at a regional scale was to be established for its usefulness when no samples are available for height measurements.

3) Examination of the existing model CANTY.

The existing stand growth model CANTY was to be examined to determine the source and magnitude of discrepancies between predictions and observations of trees in SPBL's estate. The model components of MTH, basal area, stocking and volume should be tested against the addition of various explanatory variables such as altitude, geographical type and stand age. The model can be examined quantitatively and qualitatively.

4) To develop a new model for SPBL to use at both stand and tree level.

Compared to CANTY, the new model was to be built more accurate and flexible for use at both stand and tree level. A suitable equation form would be chosen from various difference equations. The most relevant explanatory variables, if any, needed to be identified and included into a model to improve projection ability. MTH, basal area per hectare, stems per hectare, volume per hectare, and diameter distributions were to be included at a stand level, while diameter at breast height outside bark and surviving probability for each individual tree in a plot were to be studied at individual tree level. The performance of the new model was to be examined and a computer simulation program was to be developed to provide a tool to implement the model.

For projection of diameter of individual trees, an approach based on relative basal area and an approach based on sigmoid difference equations had great potential for reliable predictions. The two approaches were to be studied and compared. The output from an individual-tree model should be consistent with the one from stand level model. Adjustments were to be made when the sum of individual trees in a plot was not the same as projections based on a projection equation at stand level.

5) To develop a juvenile growth model for radiata pine in Canterbury.

To intensify the study, more data including temporary plots were to be sought. The common model components of mean top height (instead of mean height), basal area per hectare, and survival were to be included. Main environmental and silvicultural variables affecting growth were to be identified with appropriate methods and regression procedures. Juvenile growth can be simply expressed as a function of conditions of genetics, seedling status, sites, and treatments at any stage. To reflect growth responses to these factors yield-age equations were to be employed.

1.3 SYMBOLS

Throughout this thesis, unless otherwise stated, the following symbols and definitions apply.

α, β, γ or *a*, *b*, *c*: parameter or regression coefficients in an equation

Alt: altitude above sea level (m)

CanJuv: a model established in this study for juvenile trees in Canterbury

CanSPBL: a model established in this study for trees in SPBL's estate in Canterbury

CANTY: an existing model established by NZFRI for trees in Canterbury

d or *dbh* or *dbhob*: individual-tree diameter at breast height out of bark (cm)

$\bar{d}$: arithmetic mean diameter (cm)

$d_{\bar{g}}$: quadratic mean diameter (cm)

d_{gl}: individual-tree diameter at ground level outside bark (cm)

d_{max}: maximum diameter at breast height in a plot or stand (cm)

d_{min}: minimum diameter at breast height in a plot or stand (cm)

d_{std}: standard deviation of diameter in a plot or stand (cm)

d_{var}: variance of diameter in a plot or stand (sq.cm)

G: net basal area per hectare (sectional area at breast height, m^2/ha)

GF: growth and form rating used to grade genetic improvement

G_{GL}: net sectional area per hectare at ground level (m^2/ha)

GEOGRPHY: geography (it appears as graph lable)

h: total individual-tree height from the ground level (m)

H, *MTH* or $\bar{h}_{100}$: mean top height (m)

ha: hectare

MSE: mean square error obtained with regression procedures

N: number of stems per hectare (stocking)

n: total number of observations in a plot or sample size depending on context

PSP: permanent sample plot

RMS: residual mean squares (same as *MSE*)

RSS: residual sum of squares (or *SSE*: sum squares error)

S: survival of trees in a plot

SI: site index – mean top height at age 20 years

SPBL: Selwyn Plantation Board Ltd.

T: age of stand

v: individual-tree stem volume (m^3)

V: stand volume per hectare (m^3/ha)

CHAPTER 2

LITERATURE REVIEW

Two relevant components in this literature review are growth characteristics of *Pinus radiata* and modelling perspectives. Biological characteristics of *Pinus radiata* growth direct the fundamental aspects of the modelling process. A good understanding of modelling systems is a prerequisite to the establishment of appropriate models which are able to predict accurately.

2.1 GROWTH OF *PINUS RADIATA*

Growth of *Pinus radiata* at any age can generally be expressed as a function of genetics, environment, silvicultural operations and possible interactions among these factors. Genetic variety represents the difference in internal properties of the species and the environmental factors form the external world for the trees. Silvicultural operations can directly affect growth and indirectly affect growth by their impacts on microenvironments.

2.1.1 Genetics and growth of *pinus radiata*

Radiata pine seeds were introduced into New Zealand from mainland California in the period between 1850 and 1880. Genetic improvement of radiata pine in New Zealand has been conducted since the 1950s. Genetic varieties are classified as GF (growth and form), LI (long internode), DR (Dothistroma resistant) and HD (high wood density), as outlined in Genetics and Tree Improvement Research Field (1987) and Maclaren (1993). GF is a term commonly used in practice in New Zealand, and a high associated GF number implies a high growth rate in size and a high percentage of acceptable stems. Vincent and Dunstan (1989) recognised seed categories as bulk seed, felling tree selected, climbing tree selected, open-pollinated in orchard, and controlled-pollinated in orchard. A marked gain in growth and acceptable stems as crop trees has been reported for trees from controlled-pollinated seeds. The effects of genetic variation should be incorporated in growth and yield models if significant improvement can be detected.

2.1.2 Environment and growth of *Pinus radiata*

Tree growth is a result of complicated chemical processes that can be affected by various environmental components including atmosphere, light, temperature, rainfall, soil type, soil texture, nutrients, soil depth and water contents. Jackson and Gifford (1974) indicated mean volume increments of radiata pine in New Zealand were significantly related to precipitation, temperature departure from optimum, stand age, effective soil depth, and total soil nitrogen and phosphorus. Hunter and Gibson (1984) concluded site index in the North Island could be predicted from the extent of departure in pH from 6, mean annual temperature departure from 12, soil fertility, penetrometer resistance of soil, and soil depth.

Site index, the expected mean top height at age twenty for radiata pine, is mostly used to describe site quality for wood production purposes. It ranges from 15 to 40 over the whole of New Zealand and it is usually higher in the North Island than in the South Island (Burkhart and Tennent, 1977; Eyles, 1986). Altitude is often a good indicator of growth rate and it has been included in many growth models in New Zealand (Clutter and Allison, 1974; Woollons and Hayward, 1985; Mason, 1992, for example).

The earliest introduction of *Pinus radiata* to New Zealand was in Canterbury in the 1850s (Shepherd, 1990). The first planted production forestry company in New Zealand, Selwyn Plantation Board, started in 1879 and its assets were purchased by

Selwyn Plantation Board Limited (SPBL) in 1993. Radiata pine is planted mostly under 600 metres above sea level in Canterbury and its growth rate is below and close to the national average as a result of physical conditions in the region.

Geographically Canterbury has the largest plains in New Zealand (Ministry of Forestry, 1994). To the west of the plains are hills and mountains built mainly of sedimentary rocks; and to the east in mid-Canterbury are the coastal hills of Banks Peninsula built of volcanoes. Soils in Canterbury developed mainly from wind-blown dust (loess). Soils of plains (including inland alluvial basins) are usually shallow and underlain with gravel that causes low moisture storage capacity. Soils of down lands, hill country and old mountain surfaces in the high country are deeper loess-based and their water storage capacity is better. Rainfall in Canterbury is far lower than the optimum of 900 to 1800 mm for radiata pine in New Zealand suggested by Hinds and Reid (1957). Ledgard and Belton (1985) concluded precipitation accounted for an average of 75% of variation in wood production for several species and 60% of that for radiata pine. Higher rainfall in the high country combines usually with lower temperature, however, which might also affect growth. Northwest gales are risky for crops on inland plains and they result in blowing over, breakage, toppling, and resin pockets (Clifton, 1969; Mason, 1989; Studholme, 1989; Turner, 1989). Heavy snow can damage trees in some high country areas. The fact that drought, windthrow, occasional heavy snow, and periodical weather fluctuations cannot be precisely predicted may affect and restrict the performance of growth and yield models, especially for the component of surviving trees per hectare.

2.1.3 Silvicultural operations and growth of *Pinus radiata*

Various silvicultural operations such as land preparation, planting, fertilisation, weed control, thinning, pruning and harvesting all affect growth of radiata pine in various ways. Snowdon and Waring (1984) noted two types of response curve: Type 1 represents temporary effects of operations such as weed control that may reduce time to reach maximum growth rate; Type 2 represents the long-term effects of operations such as fertilisation that may change stand productivity continually over a long period.

Harvesting mature crops may affect survival and growth of trees of subsequent crops considerably through compaction and removal of nutrients. Removal of all logging residue and annual litter resulted in 12 % loss of volume production at age 16 on a fertile site in Kaingaroa Forest (Ballard and Will, 1981). Soil nutrients, affected by harvesting intensity and amount left on the floor, was closely related to diameter growth at age 5 (Dyck and Bow, 1992). Residues from harvesting can materially maintain the amount of soil moisture and reduce summer soil temperature (Farrell, 1984). Harvesting using machinery caused soil compaction leading to reduced root growth (Greacen and Sands, 1980; Sands, 1983). Murphy (1983) found that survival and growth of trees on disturbed areas of skid trails were considerably reduced. Land clearing options after harvest to facilitate site access and planting have been listed in Mason and Cullen (1986), which is useful for practising managers to plan silvicultural operations efficiently.

Such cultivation options as burning, windrowing, ripping, bedding and rolling have also been tried throughout New Zealand. Heavy and wet soils in Northland, Westland and Southland are generally those that provide the bigger gains from cultivation. Ripping and bedding have shown significant effects on growth (Hetherington and Balneaves, 1973; Williamson, 1985; Mason *et al.*, 1989). Higher growth rates are usually gained within windrows rather than between windrows since more nutrients are stored in residues within windrows (Ballard, 1978a; Farrell, 1984; Dyck *et al.*, 1989). Burning may cause nutrient loss and reduce growth, but not all burning led to decreases in growth and survival in Australia (Hall, 1985).

Weeding (killing weeds by herbicide, machinery, grazing, or manually) could improve growth of young trees in most circumstances. Madgwick (1994) summarised 26 experimental results, of which more than two thirds showed significant improvements for height and diameter from weeding. Balneaves (1982) found sustained increased height growth following the application of herbicide on droughty sites in Canterbury, New Zealand. Removal of weeds around trees is an essential option at establishment on frost-prone sites (Menzies and Chavasse, 1982; Balneaves *et al.*, 1988). Sands and Nambiar (1984) studied water relations of 28-month-old radiata pine in competition with weeds in Australia and found that severe water stress with consequent productivity loss occurred in trees with weed competition in their first growing season due to shallow root systems. By contrast, trees without weed competition were not water stressed over this period even when planted at 3 times their normal stocking rate.

Nutrient deficiency in soils could cause various problems and fertilisation is an artificial correction. Radiata pine is sensitive to nutrients due to its rapid growth rate: growth responses to nitrogen, phosphorus, potassium, boron and magnesium have been frequently documented in New Zealand, but the magnitude is site dependent. Nitrogen fertiliser could often increase growth significantly on sands and podsolised soils while phosphorus fertiliser could affect trees positively on groundwater podsols and leached clays (Shula, Hunter and Hoy, 1983). Potassium could limit growth on serpentine soils, impoverished podsols and podsolised leached sands (Ballard, 1978b; Mead and Gadgil, 1978; Will, 1985). Boron deficiency has been noted in Nelson and Canterbury (Will, 1985). Magnesium deficiency occurred on soils of deep and coarse rhyolitic tephra and treatments of dolomite fertiliser were investigated by Hunter (1996). Deficiency can be detected by visual symptoms, measured with foliage analysis, and assessed by soil types (Will, 1985; Hunter *et al.*, 1991).

Pruning may create knot-free clear wood but reduce wood production due to the loss of foliage (Shepherd, 1961; 1967; Van Larr, 1973; Lange *et al.*, 1987). A higher financial profit was made from pruned quality logs, the amount of compensation from which justified the pruning practice. Pruning strategy is determined by target diameter over stubs (DOS), required clear log length, market, costs, and growth rate of trees. Many companies in New Zealand manage a pruning height up to 6 m through several lifts but Selwyn Plantation Board Limited (SPBL) has practised 2.5 m on average through one pruning and one thinning in the last thirty years. The negative pruning effect was more severe on growth of basal area than of height (Shepherd, 1961; Shepherd, 1967; Sutton and Crowe, 1975). Tree growth was related to pruning intensity and West *et al.* (1982; 1987) included mean crown length and pruning height in the growth model EARLY to calibrate the effect.

Thinning is an efficient operation to control stocking and spacing. Thinning may affect tree and stand basal area, but thinning from below was found not to affect mean top height (Whyte and Woollons, 1990). When trees start to compete, the removal of some competitors opens up space and increases the increment of diameter of remaining trees (Lane Poole, 1944, for example). The thinning effect could be reduced by increasing time after thinning. Whyte and Woollons (1990) reported a result of a replicated thinning experiment in Kaingaroa, New Zealand and concluded that, compared to other treatments the loss of basal area from the heaviest thinning treatment at age 7 years with a residual 200 trees per hectare continued to widen up to age 24 years. Woollons and Hayward (1985) created separate equations for unthinned and thinned stands to represent the trajectory difference of basal area after thinning. Not all thinning caused differences in basal area. James (1976) found constant basal area growth over a range of stockings. The findings from such thinning trials could help forest managers to decide on the intended final crop stocking.

Improved nursery treatments, seedling handling, planting practice and some other operations could all contribute to improvements in survival, fast growth and a more uniform crop, and allow lower initial stockings at planting for selection to reach the designed quality crop, and as a result the cost of wood production would be reduced.

To avoid confusion it seemed clearer for this study to name trees younger than five years old as juvenile growth. After examining branching characteristics of crowns, Jacobs (1937) classified the natural growth of radiata pine into five stages: juvenile, adolescent, bulbous, mature, and senescent. West *et al.* (1987) described a growth model for age 4 to 14 as EARLY, particularly dealing with thinning and pruning. Mason (1992) defined a model for the phase from planting to just before thinning (at age 5) as an initial growth model. A rotation of a radiata pine crop in New Zealand usually lies between 25 and 30 years with a variation in terms of region, market, and yearly wood flow limits.

Genotype, site quality, and silvicultural operations all affect growth of radiata pine plantations. Understanding the characteristics of the species provides fundamental biological information used as a basis for modelling growth and yield.

2.2 MODELLING GROWTH AND YIELD

A growth and yield model in a forestry context is an abstraction of natural dynamics of trees, stands and whole forests. Future production of wood and growth responses to silvicultural operations can be predicted with growth modelling techniques. Forest managers use growth models to conduct production planning and researchers use growth models to analyse growth responses to silvicultural operations and environmental change.

Procedures of growth modelling include the dynamic cycle of data preparation, model construction, model validation, model implementation, and model re-calibration with a refreshed database. Model categories and some issues on the above aspects, especially on model construction and model validations, are reviewed here.

2.2.1 Categories of growth and yield models

Growth and yield models can be classified in many ways. Process models, also known as mechanistic models, usually emphasise the physiological response to environment or treatment (Bunnell, 1989; Botkin, 1993) while lower level, management-oriented models are based mainly on statistical relationships of easily measured components of trees or stands which are practically applicable to management planning. Process modelling is used to understand the mechanism of tree growth but it is difficult in practice for a manager to use such an approach (Goulding, 1994; Vanclay, 1994). A model can be deterministic or stochastic in terms of whether or not it includes a random component. A probabilistic component may be necessary to estimate parameters in some circumstances but not all of them have been found helpful in practice for decision makers to use this capability (Garcia, 1983; Garcia, 1994).

According to the level of resolution, growth models can be classified as stand level and tree level (Burkhart, 1977). Model alternatives are regarded as complementary systems rather than mutually exclusive options. Munro (1973) defined three categories of models as whole stand, distance-independent individual-tree, and distance-dependent individual-tree models. Stand models use stand values as the basic modelling unit and size class distribution may be derived in some cases. Individual-tree models use individual trees as the basic modelling unit.

2.2.1.1 Stand models

Stand models usually contain the common components of mean top height, basal area, stocking, and volume based on per unit area. Many stand models include a component of size-class distribution applied to plantations (Woollons and Hayward, 1985; Kuru *et al.*, 1992; Xu *et al.*, 1992). To apply stand models, only stand statistics are required as inputs.

The earliest applications of stand models were yield tables, which were established mainly using graphical methods and temporary plots for even-aged forests. Normal yield tables predict yields at a range of ages for fully stocked stands in a given site (Bruce, 1926). Empirical yield tables were built with random samples to estimate yields for stands of average stocking (Husch *et al.*, 1982). Empirical yield tables were a more objective approach than normal yield tables but are not sensitive to variation in stand density. The third category of yield tables was variable density yield tables which provide yield estimates for varying stand densities. Lewis (1954) for example produced a variable density yield table for unthinned radiata pine in New Zealand, but it could be successfully applied to some thinned stands because of variable density.

Stand density is the extent to which a site is used by trees. Tree size can vary substantially with stand age and site a simple measure of living stems per unit area cannot describe the site occupancy adequately. Other stand density measures are basal area, crown closure, crown competition factor (Krajicek *et al.*, 1961), stand density index (Reineke, 1933), spacing index or relative spacing (Beekhuis, 1966), tree-area ratio (Chisman and Schumacher, 1940), and point density (Bitterlich, 1947; 1948). It is desirable that a measure of stand density is independent of stand age and site quality but related to increment of stand volume (Spurr, 1952). Zhao *et al.* (1991) indicated that height-square of all trees per unit area (i.e. $N\bar{h}^2$/ha, where N is stems/ha and $\bar{h}$ is mean height) had the least correlation with age but significantly positive correlation with volume increment for unthinned larch. It was impossible to simultaneously achieve absolutely low correlation to age and absolutely high correlation to volume increment, however, given that age and site are related to volume increment.

Growth and yield equations are concise forms compared to tables and graphs. Equations can be classified as yield equations, differential equations and difference equations according to their mathematical forms. Yield equations predict cumulative yield with age and some other easily available explanatory variables. Differential equations describe the growth rate at any moment and are very useful in deriving models based on some biological assumptions (Von Bertalanffy, 1949). Difference equations project a future state from the current state and produce usually good estimations due to the close correlation between the two states. Compatibility is ensured if a derivative-integral relationship among the three forms of equations exists (Clutter, 1963; Clutter *et al.*, 1983). Using historical development and mathematical forms, Vanclay (1994) classified equations in more detail into empirical yield equations, empirical growth equations, compatible growth and yield equations, theoretical equations, and systems of equations.

Models currently used in New Zealand are mostly stand models based on a state-space system (Garcia, 1984; 1988; 1994). In a state-space system, a state is defined by several state vectors. In New Zealand, it is usually expressed in terms of mean top height, basal area per hectare, stem numbers per hectare, and sometimes crown closure. It is assumed that future states can be predicted by future management inputs and current states. The behaviour of the system is described by a transition function and an output function. The output function is a volume per hectare predicted from state variables. The system has proved to be helpful for general planning (Goulding, 1994; Garcia, 1994). It would be even more useful when planning multiple product log supplies if size-class distribution components were also included (Whyte and Woollons, 1992).

Size-class distribution models can provide information of tree frequencies of different sizes, with which log assortments can be derived. Many approaches have been in use to predict or project diameter distributions. The earliest application was based on classical stand tables and diameter increment (Husch *et al.*, 1982). Various probability density functions have shown great potential in fitting size-class distributions. These functional forms included normal, lognormal (Bliss and Reinker, 1964), Gamma (Nelson, 1964), Johnson S_B and S_{BB} (Johnson, 1949a; Johnson, 1949b; Nelson, 1964; Hafley and Schreuder, 1977), Weibull (Bailey and Dell, 1973), reverse Weibull (Kuru *et al.*, 1992; Xu *et al.*, 1992), and Beta function (Lenhart and Clutter, 1971; Maltamo *et al.*, 1995).

The three-parameter Weibull distribution function is the most used for plantations because of its flexibility, simplicity, and reliability (Bailey and Dell 1973, Whyte 1992). Techniques for estimating and projecting its parameters have been developed over time. A percentile approach can be seen in Borders (1987b). An alternative percentile interval method can be found in Koesmarno (1996). The method of moments (Ek *et al.*, 1975; Garcia, 1981) provided an easy approach to parameter estimation. The employment of maximum diameter to estimate the location parameter in a reverse Weibull function resulted in a better fit as maximum diameter is more biologically meaningful and predictably changeable with time than minimum diameter (Kuru *et al.*, 1992; Xu *et al.*, 1992). On average an improvement in reducing residual sums of squares by 5.4% was achieved by inclusion of skewness into the moments method to estimate traditional Weibull functions (Lindsay *et al.*, 1996) and the gain was mainly for some negatively skewed plots. The work from Hyink and Moser (1983) and Knoebel and Burkhart (1991) indicated that the parameter recovery technique was better than the prediction technique in the projection of future parameters. Diameter distribution characteristics, such as maximum diameter, mean diameter, diameter variance, can be projected with more confidence than the parameters in a Weibull function.

A problem of distribution models is the spatial correlation of tree diameters (Garcia, 1991). Such measures from a random plot as extreme values, variance, and percentiles vary with plot size and they are not theoretically unbiased estimates of stand. The magnitude of the impact has not been explored fully, however, and may not be excessive for many practices (Whyte and Woollons, 1992). A 3-parameter Weibull distribution is limited to the description of uni-modal shapes only. Cao and Burkhart (1984) compared mixtures of uni-modal distributions of modified Weibull cumulative distribution functions with a 3-parameter Weibull distribution, and found the former was superior in the case of thinned stands of *Pinus taeda* plantations in the United

States of America. For irregular stands, individual-tree models provide the most flexibility.

2.2.1.2 Individual-tree models

Individual-tree models use the individual tree as the basic unit to model growth of tree diameter (or basal area), height, mortality, and possibly crown characteristics. They require detailed inputs and provide detailed outputs. Though individual-tree models are more costly to build than stand-level models, they provide a useful alternative for irregular size-class distributions. Most individual-tree models describe the increment of diameter or basal area and a few models predict diameter and height based on difference equations (Zhang *et al.*, 1996a, for example).

Two classes of individual-tree model are distance-dependent individual-tree models and distance-independent individual-tree models (Munro, 1974). The distance-dependent individual-tree model requires not only the measurements of tree size but also the measurement of tree location (Daniels and Burkhart, 1975; Tennent, 1982). Distance-independent individual tree models require no spatial information of neighbours (Clutter and Allison, 1974; Clutter and Jones, 1980). The performance of distance-dependent individual-tree models lies in reliability and usefulness of competition measures that describe the reduction of tree growth. The general unavailability of inter-tree-distance in permanent sample plots in practice limits the implementation of such models. Improving predictions of individual-tree growth through the use of tree spatial information has had a mixed success in the literature (Daniels and Burkhart, 1975; Clutter *et al.*, 1983; Bruce and Wensel, 1987; Biging and Dobbertin, 1995). It was of lesser value for plantations in terms of full control of spacing. A distance-dependent individual-tree model has been established for radiata pine in New Zealand (Tennent, 1982) and diameter increment and mortality equations indicated a requirement for further improvement. Relative basal area projection from an initial individual-tree list or stand table (Clutter and Jones, 1980; Pienaar and Harrison, 1988) has been demonstrated by (Borders and Patterson, 1990) to be a method superior to Weibull distribution and percentile-based methods in projecting stand tables.

In an individual-tree model system survival probability of each tree needs to be estimated and logistic regression is most commonly employed (Hamilton and Edwards, 1976; Hann, 1980; Hamilton, 1990; Vanclay, 1991; Avila and Burkhart, 1992; Zhang *et al.*, 1997). The projected value with a logistic equation is bounded between 0 and 1. The inclusion of good predictors into models could increase prediction reliability.

Two levels of models are related. The individual-tree model should not be separated from the stand-level model and both should be integrated in a compatible manner to provide enough information for multiple levels of decision-making (Leary, 1979; Daniels and Burkhart, 1988). Stand models could be used for general and long-term planning while individual-tree models could be used for detailed planning (assorting timber for example). For even-aged plantations the stand model is most accepted in New Zealand and it is assumed that stand level estimates are less inaccurate (Garcia, 1991; Goulding, 1994). Somers and Nepal (1994) proposed an algorithm of linking an individual-tree model with a whole stand model based on the assumption that stand level components were estimated correctly and individual-tree growth should be adjusted. Ritchie and Hann (1997) evaluated the individual-tree method and the stand disaggregation method and concluded the former was superior to the latter. It was not evident, however, when the crown ratio was eliminated from the individual-tree model. Knowe *et al.* (1997) compared three methods of projecting stand tables for young red alder plantations and concluded an individual-tree modelling approach led to a better estimation than a Weibull distribution and a relative basal area projection approach. To assign survival probability for each tree or each diameter class through relative basal area only, seems inadequate in that study. It is argued that some more variables with a logistic regression procedure might improve the projection ability.

2.2.2 Data preparation and evaluation

Vanclay (1994) addressed the general data requirements for growth modelling and classified data as passive monitoring data and experimental treatment response data. Data should represent a whole population, cover the full range of temporal and spatial distributions and sample site and stand conditions widely. Some plots at extreme values of various conditions were considered more useful in building a model, in which future production would be used to interpolate rather extrapolate. For permanent sample plots, re-measurement intervals should be long enough to obtain increments that overwhelm measurement errors but short enough to relocate plots easily. All trees and plots should be identified and all measurements should use consistent and unambiguous standards.

After a database for PSP measurements has been established, stand variables of age, mean diameter, basal area, and the number of stems per hectare are usually easy to estimate and compute from a tree-level database (Spurr, 1952; McEwen, 1978; Husch *et al.*, 1982; Clutter *et al.*, 1983). Plot volume is the summation of individual tree volumes in the plot.

The strengths and weaknesses of a database should be examined and understood fully before being used. Data quality is usually displayed with tables, graphs as well as verbal text. Two crucial graphs were promoted by Vanclay *et al.* (1995) to display database characteristics for stands of plantations, one with site index versus age and another with stocking versus tree size. Garcia (1984, 1988 and 1994) used graphs of MTH versus age, basal area versus MTH, and stocking versus MTH to display data quality of state variables. Many different graphs were employed to display the data quality for some common or particular uses.

2.2.3 Common issues in growth model construction

<u>2.2.2.1 Problems and solutions with ordinary least-squares regression</u>

Ordinary regression with a least-squares approach provides a rather robust estimator and it assumes that the regression error items are independent, normally distributed, and variances are homogeneous (Neter and Wasserman, 1974). Few data sets in forestry practice could strictly satisfy these assumptions, however. The most obvious statistical problem associated with repeated measurements of permanent sample plots (PSP) is the dependence of errors in growth models (Sullivan and Clutter, 1972; West *et al.*, 1984; West, 1995). Other problems of lesser magnitude are heterogeneous variance, nonnormality of errors, and dependence among explanatory variables.

Autocorrelation in statistics defines the dependence of regression errors over time and it may exist in modelling growth with repeated permanent sample plots. The magnitude of autocorrelation can be tested with the Durbin-Watson method. Unbiased estimates of coefficients can be obtained but variances of coefficients and variances of residual errors are underestimated. In essence it creates the problem of invalidating regression tests but poses no problem for parameter estimates with serially correlated data. Various solutions to the problem have been found by employing methods of maximum likelihood regression or generalised regression and using ordinary regression with a reconditioned data set (Sullivan and Clutter, 1972; West *et al.*, 1984; West, 1995). Sullivan and Clutter (1972) investigated the autocorrelation problem and indicated no significant difference between ordinary least-square estimation and maximum likelihood estimation. When least-square methods are used to test significance of parameters, one good solution is to randomly select and employ one measurement from each sampling unit (plot) to remove dependence. West (1995) proposed fitting models with a full data set and testing models with an autocorrelation-free data set.

Multi-correlation (inter-correlation or collinearity) is characterised by a high correlation among "independent" variables. It may produce solutions from unstable singular matrices; estimates of regression coefficients may be distorted; and the important variables may not be identified. Predictive ability is not seriously affected by collinearity, however, if extrapolation is avoided. To identify and to remove possible collinearity, principal component analysis can be approached to select a set of relatively independent variables (Hunter and Gibson, 1984, for example).

The problem of serious heterogeneous variance may be overcome by conducting variable transformations or weighted regression. Some transformation may normalise error items as well. Modellers should check for the normality of residual distributions, the range of residuals, and the lack of residual bias in relation to predictions and explanatory variables (Neter and Wasserman, 1974; Vanclay, 1994).

2.2.2.2 Site quality assessment

Site is the combination of all environmental factors of climate, soil, and surrounding organisms that affect growth of trees. Many climatic, edaphic, landform and vegetative factors were found to affect growth of radiata pine significantly (Jackson and Gifford, 1974; Hunter and Gibson, 1984; Turvey and Cameron, 1986). Altitude is often a good indicator of radiata pine growth in New Zealand and it has been included in many growth models (Clutter & Allison 1974, Woollons and Hayward 1985, Mason 1992).

Site quality has been described indirectly and directly (Clutter *et al.*, 1983). Indirect methods include the use of indicator species, and the prediction of fertility from environmental variables. To describe site quality for wood production, stand volume per unit area is the direct indicator but it is difficult to measure and can be affected by stand density and silvicultural regimes. Site index is a commonly used direct method. For radiata pine, site index is the expected height of the mean of the 100 trees with the largest diameters in a hectare at age 20 (Burkhart and Tennent, 1977). Height of dominant trees is usually sensitive to variations in site quality, unaffected by thinning from below, and less affected by varying stand density than is stand volume. Site index has been included in a number of models (Clutter 1963, for example). Site index might not be a good indicator, however (Mason, 1992; Avery and Burkhart, 1994): it might lead to imprecision where high basal area is maintained rather than height and where juvenile stands are extrapolated to the age of 20 years (index age); site index often changes periodically with age because of climatic fluctuations, and varies with different species on the same site. Height is measured with less accuracy than diameter or basal area per hectare. Site index is still useful under some conditions, despite these limitations. Further research is needed in representation of site quality (Vanclay, 1994).

2.2.2.3 Approaches to improving model projection ability

Improving model projection accuracy has been a consistently interesting subject for research. Modellers have proposed and tested various approaches.

(1) A good sample data set covering target population spatially and temporally (Vanclay, 1994; Vanclay *et al.*, 1995) forms the most desirable material for model construction. Sampling error from inappropriate or very limited databases may cause more serious estimating bias than any other sources.

(2) A way of preparing data structures properly, in particular in dealing with the time increment (or growth interval) could improve projection ability. Borders *et al.* (1987a) investigated model performance of three methods of data preparation of non-overlapping growth intervals (NI), all possible growth intervals (AI), and longest intervals (LI). It was concluded that projection ability depended on the form of model employed. Projection ability concerning data structure may also be data source dependent, which means an existence of a different conclusion for a different data source (Lee, 1998, for example). A theoretical study may be needed to clarify the issue fundamentally. This current study was undertaken solely to adopt the best way of preparing a data set for further study rather than solving theoretical problems.

(3) The choice of an appropriate equation form was necessary as different equation forms behave differently (Clutter *et al.*, 1983; Woollons and Wood, 1992) and the performance of a particular model form is data-dependent.

(4) Appropriate use of a regression estimator could create a statistically stable model. Autocorrelation may exist in modelling growth with repeated permanent sample plots. West (1995) proposed fitting models with a full data set and testing models with autocorrelation-free data set when the least-square method is used. Sullivan and Clutter (1972) investigated autocorrelation problems and discovered no practical difference between ordinary least-square estimation and maximum likelihood estimation. Maximum likelihood and generalised regression estimators are theoretically robust, but the significance of their superiority to ordinary least-squares is not clear.

(5) Incorporation of significant environmental variables could improve the

projection ability of any model. A difference equation is an efficient projection form as the initial value of dependent variables accumulates the past history of growth and accounts for a great proportion of growth differences among different sites. Some environmental variables, however, might still create differences in projection performance (Woollons *et al.*, 1997, for example).

2.2.4 Modelling height-diameter relationships

Height-diameter relationships play an important role in forest mensuration systems. Stand mean height, mean top height, and mean height for each diameter class are usually calculated directly from equations of tree height (h) on diameter at breast-height outside of bark (d). Volume per hectare is an accumulation of individual tree volumes often derived from a two-dimensional function using measured d and estimated h.

Height-diameter equations passing through the origin (d=0 cm, h=1.40 m), attaining an asymptote for large values of diameter and always having a positive slope, have been recognised by a number of authors as the most appropriate form (for example, Curtis, 1967; Garman *et al.*, 1995). The Petterson equation (Schmidt, 1967) is an example of a height-diameter equation with these properties. It has been widely used in New Zealand (McEwen, 1978; Goulding, 1995), but the fit it provides for radiata pine in various regimes needs to be examined.

Early height-diameter equations were forms of linear and linearized equations (Henriksen, 1950; Myers, 1966; Curtis, 1967) which are still in use today, especially in PSP database computational systems, because coefficients can be solved explicitly and uniquely. As the power of computers developed, non-linear equations were tried because they are often more flexible than linear equations (Larsen and Hann, 1987; Wang and Hann, 1988; Arabatzis and Burkhart, 1992; Huang *et al.*, 1992; Dolph *et al.*, 1995).

Height-diameter relationships in a region vary with tree species, stand age, site characteristics, genetics, stocking, and silvicultural treatment (Trorey, 1932; Buford, 1986; Larsen and Hann, 1987; Wang and Hann, 1988; Dolph, 1989; Knowe, 1994). A regional model can be useful especially when no samples of height measurements are available.

2.2.5 Modelling growth of whole rotations with sigmoid difference equations

Stand basal area and mean top height of whole rotation growth usually display a sigmoid-shaped curve. Stems per hectare, however, showed mostly an inverse sigmoid trend. Various sigmoid equations proposed for modelling plant biomass or tree size include the log-reciprocal Schumacher model (Schumacher, 1939; Clutter, 1963), the Chapman-Richards model (Von Bertalanffy, 1949; Richards, 1959; Pienaar and Turnbull, 1973), the Weibull model (Yang *et al.*, 1978), the Hossfeld (Woollons *et al.*, 1990), and many others (logistic, mono-molecular, Gompertz, for example). Given the currently measured values, difference equations were able to project future state more precisely than yield-age equations.

Clutter *et al.* (1983) noted such desirable properties of growth equations as compatibility, consistency, path-invariance and asymptote. Compatibility of growth and yield means the property of integral-derivative relationship of growth and yield. Consistency means no change for zero elapsed time. Path-invariance ensures reaching the same future point irrespective of the numbers of projection intervals used. Asymptotic property defines the maximum yield when stand age approaches an indefinite value.

Difference equations or yield-age equations can be described as anamorphic and polymorphic types according to the nature of family of curves these equations generate (Clutter *et al.*, 1983). In an anamorphic curve family, the yield will be constantly proportional in any two curves at the same age. Polymorphic equations do not hold the proportionality relationship. When a polymorphic equation produces curves that do not cross within the age range of interest, it is known as a polymorphic-disjoint equation. If any two curves created with a polymorphic equation can intersect within the age range of interest, the equation is then a polymorphic-nondisjoint equation.

An important model component is stems per hectare of surviving trees. Some difference equations have been described by Clutter *et al.* (1983). Serious mortality might be caused by drought, windthrow, occasional heavy snow, animals, diseases, and competition from neighbouring trees. It is clear that many of these factors are almost unpredictable. Surviving trees per hectare is a component that can never be projected exactly (Glover and Hool, 1979). To improve prediction, modellers have tried to incorporate environmental variables into the mortality model (Mason 1992, for example) and to use a variant of basic difference equations (Woollons and Hayward 1985). Woollons (1998) used two-step regression to solve conflict between data that exhibit no mortality in a short period within a small plot and equations that assume the occurrence of loss of some trees in a period. The method of assigning a value of 1 for tree loss which occurred in a plot and 0 for no tree loss does not account for the difference in the number of lost trees, but accounts for the difference between one tree loss and no tree loss over a period. The restriction of a constant interval of re-measurement can limit the use of the procedure.

Both whole stand models and individual-tree models should be established and integrated to provide enough information for multiple levels of decision-making (Leary 1979, Daniels and Burkhart 1988).

2.2.6 Modelling juvenile growth

Juvenile growth modelling for plantations of *Pinus radiata* concentrates on the state from immediately after planting to around five years just before thinning (Mason, 1992). Juvenile growth is sensitive to environment and plantation establishment practice and managers are interested in representing the effects of establishment practices within their growth models. Juvenile growth modelling needs to identify the main variables and predict growth responses to different sites and different silvicultural operations.

The objectives of plantation establishment are to achieve high survival, consistent and large tree size, high growth rate, and good tree form. Models at a stand level often include model components of survival, mean height or mean top height, diameter or sectional area at ground level or at breast height (Belli, 1987; Belli and Ek, 1988; Mason, 1992; Latif *et al.*, 1996; Leary *et al.*, 1996; Zhao and Mason, 1996; Mason and Whyte, 1997). Individual-tree models often focus on increment of height, increment of diameter or increment of basal area (Zhang *et al.*, 1996b; Nystrom and Kexi, 1997). Diameter measurements can be taken at ground level during the first several years and at breast height later when trees grow above breast height. Modelling for juvenile growth demands a choice between diameter or sectional area at ground level, and dbh or basal area, or both. To provide compatibility with older growth, dbh or basal area is needed. Usually no suitable individual tree volume equation is available for such young trees. Tree form has been rarely modelled due to unavailability of necessary measurements. Yield-age equations have been employed by most modellers (Belli, 1987; Belli & Ek, 1988; Mason 1992, 1996; Zhao and Mason, 1996) to reflect the growth response fully for different operations and site conditions. Difference equation forms can hide effects behind the initial time (T_1) and initial growth (Y_1).

Juvenile growth and yield at any stage can be expressed as a function of conditions of sites, status of seedlings, treatments and competition forces from various weeds and from trees themselves when the crown starts closure. Seedling quality can be described physically and morphologically and it can be changed by genetic improvement, nursery regimes and stock handling (Genetics and Tree Improvement Research Field, 1987; Menzies, 1988). Site index might be derived with serious bias from juvenile stands that are far less than twenty years of age. Environmental factors of climate and soil, or mapping factors of altitude, latitude, and aspects were often sought to represent site differences. Hunter and Gibson (1984) demonstrated a close correlation between environmental variables and site index. The micro-environment can be changed by such field management options as land clearing, cultivation, fertilisation, and weed management (Balneaves, 1982; Hunter *et al.*, 1991; Mason *et al.*, 1995; Mason *et al.*, 1996).

Some equation forms for early growth and yield of tree height, diameter, and survival have been proposed and used (Belli, 1987; Belli & Ek, 1988; Mason, 1992) and they showed desirable properties and fitting performances.

The issue of the relationship between juvenile growth and older growth models has arisen since juvenile growth models have been created. Mason *et al.* (1997) displayed the prediction performance and difference between IGM (a juvenile growth model) and PPM88 (a model for older growth) with data from one forest. Linking the two models smoothly may be useful but may be complicated due to the difference in data sources, model forms and variables included in the models.

2.2.7 Model validation

Model examination or model validation is a procedure that answers how closely the model's behaviour fits the real world and to what extent logically and biologically the model agrees with actual forest growth. The validity of a model is related to objectives of forest management and can only be evaluated relatively in terms of its empirical and abstract features (Goulding, 1979; Reynolds *et al.*, 1981; Soares *et al.*, 1995).

Many aspects should be considered in model validation and they were detailed by Soares *et al.* (1995) and Vanclay (1994). A model can be examined qualitatively and quantitatively. No single criterion can capture all aspects of model validity. Qualitative validation examines the biological aspects of each single module and the logical structure of the model as a whole. Quantitative validation assesses model behaviour with a specified set of data, and this procedure involves analysing prediction errors with graphical display and statistical tests. Details are described in following sections.

<u>2.2.7.1 Data requirements</u>

Data used for validating a model are regarded as observations of reality so they should be fully representative of a population. It is desirable to test models with full data covering the widest stand conditions (Vanclay, 1994). If data are sampled with severe bias, a good model may be proved wrong. Data requirements, however, may depend on the objective or scale of model validation in some sense. Data covering all stand conditions in the estate of a forestry company are required to test the adequacy of a model to trees of that particular agent. Data covering a whole country are required to test the adequacy of the model for national uses.

A convincing test can be carried out by using a set of data that is independent of the data used for building the model. To obtain a set of independent data, different approaches have been adopted for different cases by different modellers. Cross-validation (Snee, 1977) splits data into two parts: one for building the model and the other for validating the model. Vanclay (1994) proposed using data collected by an independent agent. Time is often used in partitioning data. Unbalanced data for both building a model and examining a model might be derived, however, when data are partitioned on a particular year because some variables, such as genetic improvement rating of radiata pine in New Zealand are time-dependent. A half-and-half partitioning of data for building and examining a model might be used (Snee, 1977), but most modellers tend to use more data in building a model and less in validating the model in order to increase model reliability, especially when data are insufficient.

Long-term repeated measurements provide a good base for model validation. Plot measurements with shorter periods have considerable variability and growth pattern cannot be compared (Goulding, 1979).

2.2.7.2 Model validation of qualitative properties

A model and its components should be structured logically and be biologically realistic. Agreement should be reached on the common theories of growth modelling (Clutter *et al.*, 1983; Bossel, 1991; Vanclay, 1994). The equation forms for growth and yield should own the desirable properties suggested by Clutter *et al.* (1983). The estimated parameter values and signs should agree with normal understanding of growth processes. Model sensitivity to a production function was analysed by Liu *et al.* (1989) to examine the magnitude of each component and can be carried out by examining the effect of predictions on changes in a parameter. It is a useful procedure but may be tedious to undertake when many parameters are included.

Approaches to parameter estimation should comply with the theories of statistics. If the assumptions of linear regression are severely violated for certain data available, some transformation or some other estimation methods should be used (West *et al.*, 1984; West, 1995).

When qualitative evaluation is complete, models can then be validated quantitatively.

2.2.7.3 Quantitative validation by analysing prediction errors

Quantitative examination should include: a) graphical analysis of residuals against predictions and against other variables which possibly reflect stand conditions; b) statistical regression of residuals against predictions or observations and against possible stand condition variables; c) conditional statistical tests for the model.

Inspection of graphical plots of residuals versus predictions or potential explanatory variables is an efficient way to detect possible correlation. Residual errors may display certain trends along with initial conditions or projection length if the model is biased. Mayer and Butler (1993) proposed that scatter plots of observed versus predicted values at the reference line with intercept zero and slope one is preferable although it is similar to the scatter plot of residuals versus predicted value. Linear regression of residuals on all stand variables is a statistically efficient way to detect linear correlation. The magnitude of the relationship can be clearly identified with this procedure (Soares *et al.*, 1995).

Statistical testing is important when determining the acceptance or rejection of model validity. Freese (1960) proposed three elements in examining the accuracy of a new technique or new model: the accuracy required by users, the accuracy attained in trials and whether or not the accuracy attained meets the accuracy required. A Chi-square test was proposed in his study. Reynolds *et al.* (1984) explained the underlying assumptions of Freese's test. Critical errors (the maximum error for the acceptance of the model and minimum error for rejection of the model) were recommended if the distribution of residual errors is normal with mean value of zero. Gregoire and Reynolds (1988) displayed furthermore that the test was not robust when residuals were non-normal and a misleading interpretation could result from its use. Linear regression of residuals on predicted values can be tested with a simultaneous F-test for zero intercept and slope.

Several common and simple measures used to describe model performance (Equations 2.1 to 2.4) are: average model bias, mean absolute errors, percentage mean absolute errors (Schaeffer, 1980) and model efficiency (Loague and Green, 1991).

a) average model bias $$AMB = \frac{1}{n}\Sigma(Y_i - \hat{Y}_i) \qquad (2.1)$$

b) mean absolute errors $$MAE = \frac{1}{n}\Sigma|Y_i - \hat{Y}_i| \qquad (2.2)$$

c) percentage mean absolute errors $$PMAE = \frac{1}{n}\Sigma\frac{|Y_i - \hat{Y}_i|}{|Y_i|} \qquad (2.3)$$

d) model efficiency or efficiency factor $$EF = 1 - \frac{\Sigma(Y_i - \hat{Y}_i)^2}{\Sigma(Y_i - \overline{Y}_i)^2} \qquad (2.4)$$

where Y_i = observed value, $\hat{Y}_i$ = predicted value, n = sample size

Average model bias is an average of errors for overall predictions. A value of zero AMB, an equity of the sum of positive residuals and the sum of negative residuals, is an ideal expectation but does not guarantee lack of bias when a linear or non-linear relationship between residuals and predicted (or observed) values actually exists. The positive and negative deviations may cancel each other out. Mean absolute error is an average distance of residuals from zero. A small value of MAE is preferred and it may imply different reliability when prediction scale and value vary. A relative value (percentage error) is an alternative measure but a potential problem is the heavy influence from very low observed values. It tends towards infinity as the observed value (Y_i) closes to zero. EF is a measure of overall performance and was recommended by Loague and Green (1991) and Mayer and Butler (1993). A high value of EF up to one is expected for a perfect fit, zero means a prediction equivalent to the mean value Y, and a negative value indicates a worse prediction. EF is an extension of R^2, which is the proportion of variation explained by the fitted regression line with the data. R^2 has a lower bound of zero and upper bound of one and it is not a good indicator of model performance when outliers appear or a non-linear relationship exists

(Weisberg, 1985; Vanclay, 1994). A high value of R^2 cannot guarantee a good model as repeated measurements are usually closely correlated especially when difference equations are used.

A measure to detect errors from sampling and model structure is the Janus quotient, proposed by (Gadd and Wold, 1964) and recommended by Power (1993). It is the ratio of mean squares of residuals from model validation to mean squares of residuals from model creation. A value of one indicates a good model structure and good predictive ability outside the sample in building the model. The independence and normality of residual errors also should be tested (Power, 1993).

A wide variety of measures and approaches to model validation is thus available, but none of them can individually replace all the others. It is desirable to use several tests to assess different aspects of model behaviour. It is also desirable to partition data in terms of biological differences and to examine model performance in each of several strata (Mayer and Butler, 1993; Soares *et al.*, 1994).

First of all, a suitable database for model validation and model creation needed to be established.

CHAPTER 3

ESTABLISHMENT AND ASSESSMENT OF A PSP DATABASE FOR GROWTH MODELLING

3.1 INTRODUCTION

The study reported here began with database design, data entry, storage and processing. All variables needed for plots and stands were computed and derived from the database. The design of a relational database system to complete the task is emphasised in this chapter. The properties of the main variables for growth modelling were assessed and described as well.

The data used in the study were mainly from Permanent Sample Plots (PSP) in the estate of Selwyn Plantation Board Limited (SPBL). Approximately 10 000 ha of stocked plantations are owned by SPBL (SPBL, 1998) and 90% of the stands are pure, even-aged radiata pine. All SPBL's plantations are scattered between the Waimakariri and the Rakaia Rivers in mid-Canterbury (Figure 3.1). They cover plains, hills and coastal sands. Altitudes of the PSPs ranged from 2 to 600 metres above sea level. Away from the sea, soils change from coastal sands, shallow and dry flood plains soil, to deep and wet loess hill soil. Rainfall increases with altitude and distance from coast. In each plot diameter at breast height outside bark (dbhob) of all trees was measured but heights of only eight trees (seven in a few cases) were sampled.

A continuous forest inventory system comprising solely permanent sample plots was established by SPBL during the 1960's and re-measurements were regularly taken thereafter at three to five year intervals until harvest. A few new plots were often established due to increased age or due to changes in stand conditions. The inventory was based on a sampling intensity of one per cent. These plot measurements were used for estimation of population variables, forest valuation, growth modelling, and preparation of harvest plans.

Several distinguishing features should be understood before establishing a database and using the database for growth modelling.

(a) Forests in such a typical forestry enterprise management system as SPBL's are usually divided into hierarchical layers comprising forests (or plantations), compartments, stands, plots, and trees. A forest contains many compartments, a compartment contains many stands, a stand contains many plots and a plot contains a number of trees. The levels may vary according to the scale of an enterprise.

(b) PSP plots in SPBL were established after thinning and pruning, between mostly ages seven and nine, and no further tending treatments were applied after thinning. The company has managed its plantations retaining more stocking and green crown than has been traditional in New Zealand. Pruning height varied between 0 to 6 m but was mostly around 2.5 m. Remaining stockings just after thinning were almost all greater than 500 stems/ha. The strategy is aimed at minimising wind damage.

(c) The plantations in SPBL are fully scattered between two rivers, and they are small in area but many in locations (Table 3.1 and Fig. 3.1). What this means is that those plantations represented a good sample of mid-Canterbury. The locations are independent of sites on which the model CANTY was based.

The establishment of a database for PSP measurements using a Data Base Management System (DBMS) is described in the following section.

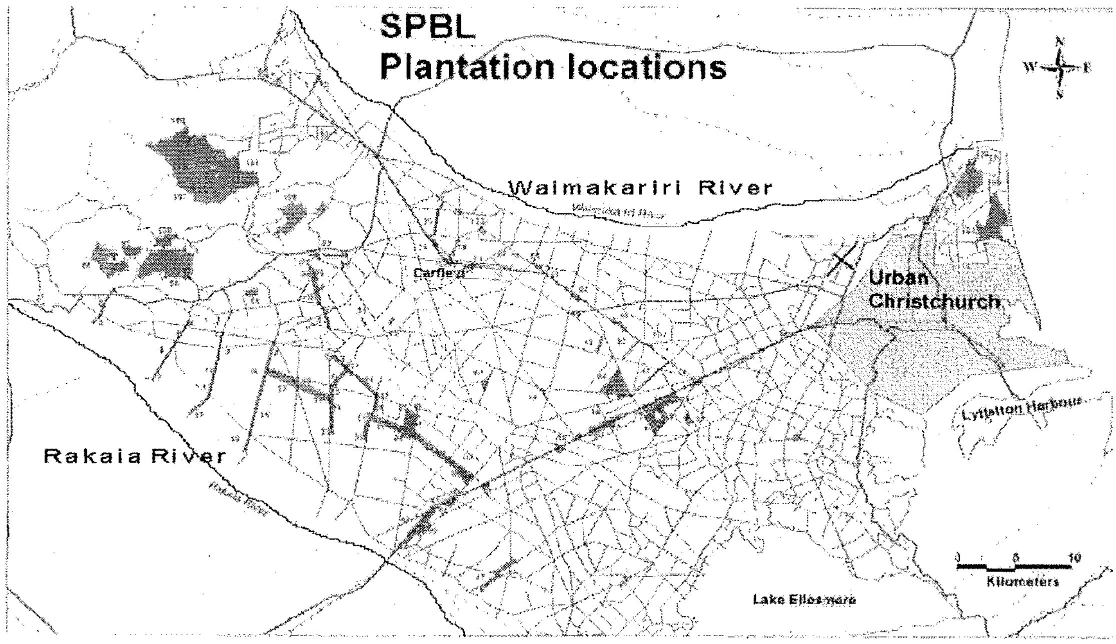


Figure 3.1: Map of SPBL plantation locations (cited from SPBL's map system)

Table 3.1: The plantation locations in SPBL and locations covered by CANTY

Source	Locations
CANTY	Eyrewell, Balmoral, Burnham, Bottle Lake, Ashley, Hanmer, Omihi, Waimate, Geraldine
SPBL	ATC Bankside, Annat, Ardlui, Bates, Bells, Bennetts, Bottle Lake, Boundary Rd, Britcliffs, Chamberlains, Chaneys, Coalgate, Collisons, Conways, Daltons, Derretts, Doyles, Earlys, Englands, Gardiners, Gillanders, Griggs, Guineys, Halls Dunsandel & Sleemans Rd, Harts, Hendersons, Holmes, Joblins, Kensingtons, Kibblewhites, Kimberley, Lighthouse, Logans, Lowmount, Magans, Mastertons, McHughs, McIntoshes, Mead, Mitchells, Mugfords, Muirheads, Mulligans, New Park, Norwood, Nursery/HQ, Olivers, Pages, Parkins, Prestidges, Ridgens, Sandy Knolls, Shellocks, Shemmings, Skellerups, Sleemans Rd, Smiths, Sollitts, Southbridge, Steeles Road, Te Pirita, Thwaites, Tilsons, Tranters, Turnbulls, Tweeds, Wattle, Webbs, Westenras, Wrights

3.2 ESTABLISHMENT OF A DATABASE FOR PSP MEASUREMENTS USING DBMS

It was necessary to establish a database for PSP measurements because the majority of records from SPBL were in hard copy (paper copy) form when the study began. A few records in recent years were stored as individual Quattro Pro files (electronic form) for each stand measurement. Specialised forestry database software for PSP measurements was unavailable for SPBL and for this study.

Data Base Management Systems (DBMS) have advantages over file-based spreadsheet systems for PSP measurements. Data separation and isolation, data duplication, and data dependence are the main limitations of a file-based system. Arnold (1994), for example, met difficulties and was unable to use spreadsheet systems to draw plot and stand information directly from a Geographic Information System (GIS). DBMS systems provide good data accessibility, security, integrity, and re-producibility. DBMS systems lead to efficiencies in data entering, storage (size), and query making (Connolly *et al.*, 1996).

3.2.1 Design of a relational database

Among several types of databases, a relational database is the most popular and useful model (Codd, 1970; Codd, 1990; Connolly *et al.*, 1996). In designing an Entity-Relation (ER) database for PSP records, the following aspects were considered.

(1) Multi-level tables for a forest enterprise's management system should include typically the hierarchical tables of forest (or plantation), compartment, stand, plot, and tree. Lower-level tables contain the summary information in higher-level tables. A forest table or compartment table could be easily derived from a stand table. More or fewer levels might be needed according to an enterprise's scale.

(2) Every attribute belonging to a particular table should be included in that table and primary keys should be identified. The forest table in this database for PSP records

in SPBL included forest name (location or region, as primary key), area and rainfall. The compartment table included compartment name (code, as primary key), forest name (foreign key), location and area. The stand table included stand name (primary key), area, geography type, main species, planting date, initial stocking, GF rating, compartment name (foreign key), and forest name (foreign key). The plot table included plot-number (primary key), area, establishing date, altitude, latitude, longitude, slope, aspect, stand name (foreign key), compartment name (foreign key), and forest name (foreign key). The tree table included tree-number (primary key), tree species, plot number (foreign key), stand name (foreign key), compartment name (foreign key), and forest name (foreign key). A record in a table could be identified by combining a primary key with a foreign key or keys. All variables listed in every table at this step were static attributes.

(3) Measurement tables store such dynamic attributes as measurement date and measured values. A tree number or plot number can appear more than one time when measurements have been taken more than once so the measurement date should be listed in measurement tables as the primary key to identify different measurements. Three measurement tables for stand level, plot level and tree level were designed. A big tree measurement table, for example, could include measurement date (primary key), diameter, diameter type (dbh or diameter at ground level), total tree height, tree status comments, possible crown measures, and all levels of foreign keys of tree number, plot number, stand name, compartment name and forest name.

(4) Treatment tables should be designed so that the level at which the treatment applied is identified and the key in the table is defined to link with other tables. In the pruning table, for example, pruning height was recorded for every individual tree in SPBL so the table included a primary key of measurement date, and foreign keys of tree number, plot number, stand name, compartment name, and forest name. The thinning table, however, included only remaining stocking and other attributes at a plot level so a primary key of measurement date, and foreign keys of plot number, stand name, compartment name, and forest name were included.

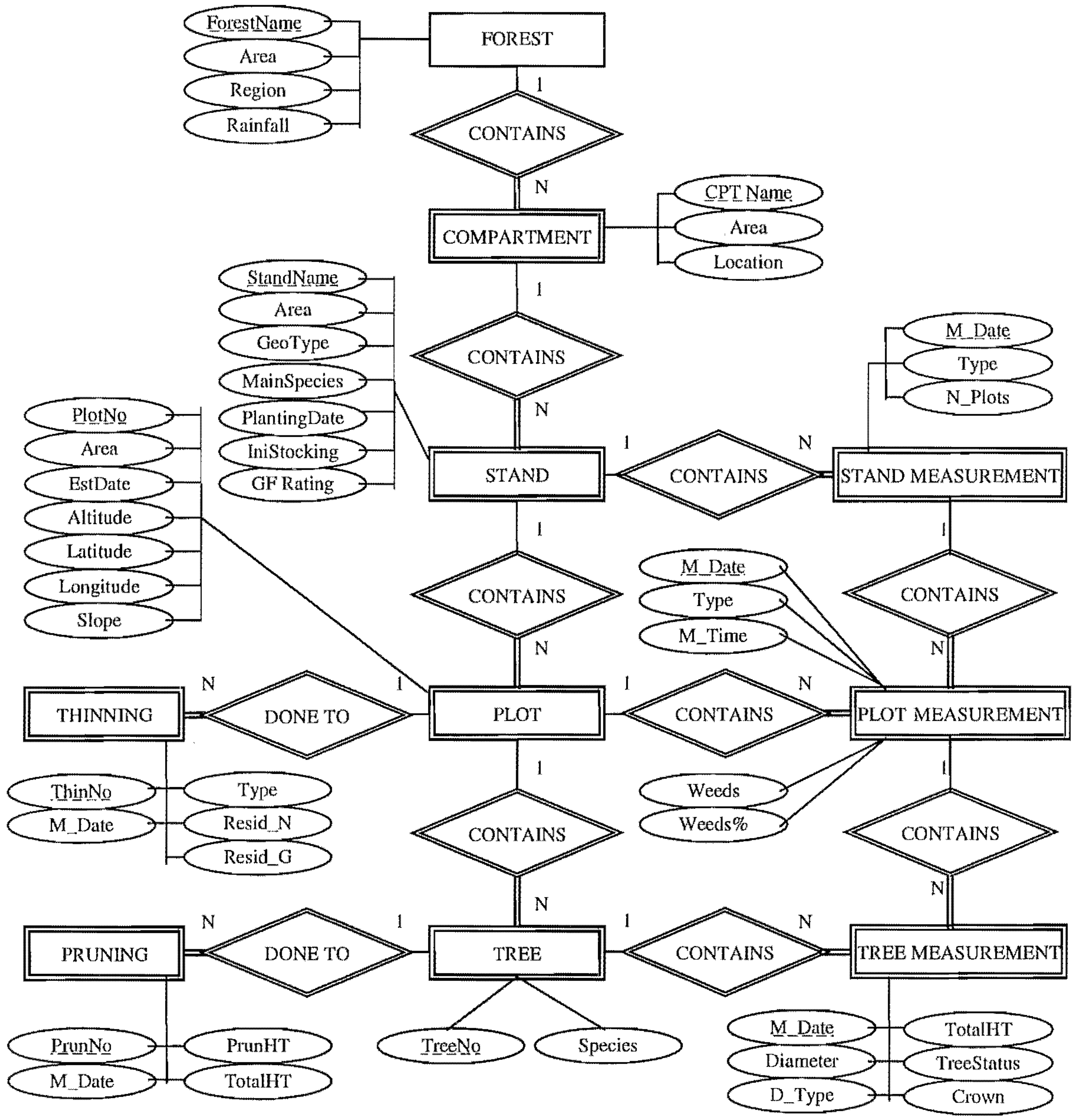


Figure 3.2: PSP database ER diagram

A PSP database Entity-Relationship (ER) diagram was shown in Figure 3.2 to display the database structure, tables and the relationship among tables. The symbols in the diagram are standard types for ER databases and they are consistent with the notation listed by Connolly *et al.*, (1996). Table structures of the database could be described as follows (key attributes are indicated with underline).

FOREST

FOR_NAME	REGION	AREA	RAINFALL

COMPARTMENT

CPT_NAME	AREA	LOCATION	FOR_NAME

STAND

STD_NAME	AREA	GEOG_TYPE	MAIN_SPECIES	PLT_DATE

INI_STKG	GF_RATING	FOR_NAME	CPT_NAME

STAND_MEASUREMENT

STD_NAME	M_DATE	TYPE	N_PLOTS	FOR_NAME	CPT_NAME

PLOT

PLOT_NO	AREA	EST_DATE	ALT	LAT	LONGI	SLOPE	ASPECT

FOR_NAME	CPT_NAME	STD_NAME

PLOT_MEASUREMENT

PLOT_NO	M_DATE	TYPE	WEEDS	WEED %	M_TIME

FOR_NAME	CPT_NAME	STD_NAME

TREE

TREE_NO	SPECIES	FOR_NAME	CPT_NAME	STD_NAME	PLOT_NO

TREE_MEASUREMENT

TREE_NO	M_DATE	DIAMETER	D_TYPE	TOTAL_HT	CROWN

TREE_STATUS	FOR_NAME	CPT_NAME	STD_NAME	PLOT_NO

THINNING

THIN_NO	M_DATE	TYPE	RESI_N	RESI_G

FOR_NAME	CPT_NAME	STD_NAME	PLOT_NO

PRUNING

PRUN_NO	M_DATE	TOTAL_HT	PRUN_HT

FOR_NAME	CPT_NAME	STD_NAME	PLOT_NO	TREE_NO

Where:

FOR_NAME:	forest name
CPT_NAME:	compartment name
STD_NAME:	name of a stand
PLOT_NO:	plot number
TREE_NO:	tree number
PLT_DATE:	planting date
M_DATE:	measurement date
N_PLOTS:	the number of plots in a stand
INI_STKG:	initial stocking
EST_DATE:	plot-establishing date
ALT:	altitude
LAT:	latitude
LONGI:	longitude
GEOG_TYPE:	geographical type - plain, hills and coastal sands
M_TIME:	measurement time – once, twice, three times …
D_TYPE:	diameter type (diameter at breast height or diameter at ground level)
RESI_N:	residual stocking after thinning
RESI_G:	residual basal area after thinning

3.2.2 Practical data manipulation

Based on the above database structure, a data entering form joining related tables was used in this study to enter data into tables simultaneously. After the completion of data entering, potential errors were checked by graphing tree height against tree diameter and tree diameter against tree age. Apparently abnormal records were checked with original sheets and identifiable errors were corrected.

PSP information was stored in a database comprising varying levels of tables, from which any new table could be derived when needed. As a result, there were 93 429 repeated measurements records in the tree measurement table, 3 500 records in the plot measurement table, and 529 in the stand measurement table. Backup files were created for safety reasons. In dBase file format, the storage size for all the raw data was about 20 megabytes, which was far less than the total size of those file-based spreadsheet files in SPBL that housed records of only the most recent measurements.

Summary statistics including growth-modelling information for each plot and stand were derived from data transaction procedures within the newly built database. The computational formulations and estimations for plots and stands statistics are described in the following section. Most data manipulation in this study was conducted with Microsoft FoxPro for its applicability to the study and its features of being able to execute as many sets of Structured Query Language (SQL) queries as needed to complete series data transactions with a single file of programming. The selection of software is a matter of users' preference; the really relevant and most important thing is to use correct algorithms to obtain needed information accurately.

3.3 Computation and Estimation of Stand Variables

Some main variables that can be derived from a PSP database to describe quantitatively a plot or stand of pure even-aged plantations are usually age, mean diameter, maximum diameter, minimum diameter, variance of diameter, mean top diameter, mean height, mean top height, stems per hectare, basal area per hectare, volume per hectare. Most of these can be calculated and estimated directly (Husch *et al.*, 1982; Clutter *et al.*, 1983; McEwen, 1978). A pre-determined standard tree volume equation for radiata pine, estimating tree volume with dbh and height, was available for the study to obtain volume per unit area with aggregation of tree volumes. A pre-determined equation form describing height-diameter relationships was also needed to estimate individual-tree height and mean top height. The next chapter (Chapter 4) describes the development of a height-diameter model with the newly built database, while all other computations of variables for plots and stands are described in this chapter.

Age

The age of plantations is usually taken from the year of planting. In this study it was obtained by subtracting the year of planting from the year of measurement with an adjustment for monthly difference. All trees in SPBL were planted during the winter. If a measurement was not taken in the months from April to July when growth of radiata pine nearly stops in Canterbury, an adjustment was made as listed in Table 3.2. The adjusted values were all the same as those listed in McEwen (1978) except that a little modification from 0.5 to 0.45 was applied to November to achieve a more smooth transition. It is important to adjust age when time intervals are not exactly a year apart.

Table 3.2: Adjustment of age for measurements taken in different months

Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Adjust value	-0.3	-0.2	-0.1	0	0	0	0	0.1	0.2	0.3	0.45	0.6

It would probably be better to assign the adjustment values for height modelling differently from the values for basal area modelling. Jackson (1976) and Tennent (1986), for example, indicated that monthly height increment as a percentage of total annual increment peaked in spring and early summer (October to December) but diameter maximum was in summer to autumn (Nov.-March). The Canterbury region was not covered by these studies, however.

Stems per hectare

Stems per hectare is the number of living trees in a plot divided by the plot area in ha. Estimates of stems per hectare for a stand are the averages of plot values, which are unbiased estimates for stands. Stems per hectare is used to describe stand density and reflect the dynamics of mortality with age.

Mean Diameter ($\mathbf{d}_{\bar{g}}$)

Mean diameter often implies in forestry the quadratic mean diameter and can be calculated with:

$$d_{\bar{g}} = \sqrt{\Sigma d_i^2 / n} \tag{3.1}$$

where $\mathbf{d}_{\bar{g}}$ is quadratic mean diameter (cm), d_i is diameter (cm) of the ith tree in a plot or stand, n is the number of trees in the plot or the stand.

Diameter Variance (d_{var})

The plot diameter variance can be obtained with

$$d_{var}(plot) = \frac{\Sigma d_i^2 - (\Sigma d_i)^2 / n}{n} = d_g^2 - \bar{d}^2 \quad (3.2)$$

where $d_{var}(plot)$ is diameter variance in a plot (in cm^2), n is the number of trees in the plot, $\bar{d}$ is arithmetic mean diameter ($\bar{d} = \Sigma d_i / n$). Arithmetic mean diameter is always smaller than quadratic mean diameter as long as tree diameters are not all the same.

The stand variance can be estimated with a cluster sampling estimation that was listed by Garcia (1991):

$$d_{var} = [\tfrac{1}{m\bar{n}} \Sigma\Sigma d_{pi}^2 + \tfrac{1-m/M}{m\bar{n}(m-1)} n_p \bar{d}_p^2 - \tfrac{1-1/M}{m\bar{n}^2(m-1)b} (\Sigma n_p \bar{d}_p)^2] / (1 - 1/\bar{n}M) \quad (3.3)$$

where d_{pi} is i^{th} dbh in p^{th} plot, m is sample size (the number of plots sampled in a stand), M is population of plots (the number of all plots in a stand), n_p is the number of trees in p^{th} plot, and $\bar{n}$ is average of number of trees.

Basal area per hectare (G)

Plot basal area is the sum of individual tree basal areas in a plot. The plot basal area is

$$G_p = \pi \Sigma d_i^2 / 40000 = 0.00007854 \ \Sigma d_i^2 \quad (3.4)$$

where G_p is plot basal area (m^2), d_i is diameter (cm) of the i^{th} tree in the plot.

Plot basal area divided by plot area (in ha) is the basal area per hectare. Stand basal area is an average of plot basal areas expressed on per hectare basis.

Maximum Diameter (d_{max}) and Minimum Diameter (d_{min})

The maximum diameter in a plot is the largest diameter within the plot and the minimum diameter the smallest.

$$d_{max} = \max(d_i) \quad (3.5)$$

$$d_{min} = \min(d_i) \quad (3.6)$$

The maximum diameter in a stand is estimated with the largest diameter within the stand; it is the largest among the maximum diameters of plots. The minimum diameter in a stand is the smallest diameter among the minimum diameters of plots. The estimates for both maximum and minimum diameters of stands are biased. Maximum diameter in a stand can be underestimated with plot value and minimum diameter in a stand can be overestimated (Garcia, 1991).

Mean Top Diameter

Mean top diameter is defined in New Zealand as the quadratic mean diameter of 100 trees with the largest diameters in a one-hectare area (Burkhart and Tennent, 1977; McEwen, 1978; Goulding, 1995). Mean top diameter can be calculated using the same formula as mean diameter but only top trees are included. The number of top trees in a plot was determined proportionally to the plot area. Four largest diameters should be chosen for plots sized 0.04 ha and two chosen for 0.02 ha plots.

Mean Top Height (H, MTH or $\overline{h}_{100}$)

MTH is the corresponding height to the mean top diameter estimated with a height-diameter equation. A Petterson equation with exponent -5.0 was adopted for the new database after a study of height-diameter modelling described in the next chapter.

Mean Height (MH or $\overline{h}$)

MH is the corresponding height to mean diameter estimated with a height-diameter equation. Individual tree height in a stand can be estimated with tree diameter.

Volume

Plot volume is the sum of individual tree volumes. Plot volume divided by plot area (in ha) is the volume per hectare. Stand volume can then be estimated with plot volume.

Three forms of tree-volume equations were identified (Spurr 1952, Husch 1982): local tree-volume equations express tree-volume as a function of tree diameter at breast height only; standard tree-volume equations estimate tree volume in terms of diameter at breast height and total height; and form class tree-volume equations estimate tree volume with diameter, height and a form measure. For this study, no volume measurements were available to build a new tree-volume equation, so an existing standard tree-volume equation for Canterbury (New Zealand Forest Research Institute, 1991), the same form as listed in McEwen (1978), was adopted.

$$v = \exp\{\alpha \ln(d) + \beta \ln[h^2 /(h-1.4)] + \gamma\} \tag{3.7}$$

where v is tree-volume in m^3, d is dbhob in cm, h is tree height in m estimated with height-diameter curve, α, β, and γ are coefficients.

Mean pruning height

The pruning height for every tree was available in SPBL's PSP records. Mean pruning height in a plot is the arithmetic mean height of all trees. The pruned height of each tree usually changed little within any plot.

Such site variables as altitude, latitude, rainfall and geographical type can be retrieved directly from plot, stand or forest tables. Many other variables such as weight, form and surface area are not obtainable from the database. The properties and quality of main variables from a viewpoint of growth modelling are described in the next section.

3.4 Description of Main Variables for Growth Modelling

The strengths and weaknesses of a database in terms of growth modelling should be examined and understood fully before being used. A reliable model can be built only with data of high quality. Tables and graphs should be used efficiently to display the quality of PSP data for growth modelling. Two crucial graphs were suggested by Vanclay *et al.* (1995) to display database characteristics for stands of plantations - one with site index versus age and another with stocking versus tree size. Garcia (1984, 1988, 1995) used graphs of MTH versus age, basal area versus MTH, and stocking versus MTH to display data quality for his state variables.

All observations at plot and stand levels were imported to SAS files from the database file format in the study. Data were screened and then partitioned for model building and model examination. The main variables were described in a tabular and graphical manner. Four graphs, site index versus age, mean top height versus age, basal area/ha versus age, and stocking/ha versus age (instead of tree size) were used to display the development of main variables with age. This was in accordance with model equations, in which age would be a major independent variable for projecting each state variable.

In reality a model may be used at a plot or stand level. The whole stand model can be created with either plot information (the case for most models) or stand level information (estimation from plots within a stand). It was decided to use plot statistics to build a new model for the following reasons.

1) When preparing stand level statistics, some newly established plot measurements or plot measurements with one or two measurements missing would have to be dropped in order to keep a stand with consistently the same plots in all historical re-measurements. These plot measurements might be all kept, however, when plot statistics were employed to build a model.
2) Using plot level statistics could minimise the loss of information located in individual plots. One plot with high stocking and the other with low stocking could be modelled with possibly different paths, which could be merged by averaging two plots as stand estimates. Single species, even-aged plantations might be expected to be uniform in density but in reality variations are very high in some stands, as revealed in the chapter on forest sampling analysis.
3) A model created with plot information was assumed of being capable to work well at stand level. Least-squares procedures can result in an averaging effect from plots.

Data were screened and the following plots were identified and removed:
1) fertilised plots on coastal sandy soils;
2) plots with heavy mortality caused by rare heavy snow and windblow, which implied at least six trees dead annually in each 0.04 ha plot and three in each 0.02 ha plot;
3) plots with fewer than two measurements over time.

On average, there were 6 plots per stand. Time intervals of re-measurement averaged three years. 95% of plots were 0.04 ha in area while the remainder were 0.02 ha.

The database was partitioned for model building and model examination. Altogether 2996 plot measurements were available: 505 plot measurements were chosen for model validation at a later stage and 2491 for the main model building. Table 3.3 lists the summary of plot variables in the data set for model building and Table 3.4 for model validation. Figure 3.3 displays the plot development pattern of main components of site index, mean top height (MTH), basal area (G), and stems per hectare (N) with data for model building and Figure 3.4 with data for model validation.

3.4.1 Data for model building

There were 2491 plot measurements within 950 plots for model building. A summary of plot variable estimates with a breakdown for plains and foothills is listed in Table 3.3. Figure 3.3 displays the growth patterns of the main variables. This indicated the following characteristics.

(1) Stand age ranged from 8 to 29 years old. No measurement was taken before 7 years old and after 30 years old. A denser frequency from 8 to 16 years of age and a less dense frequency 25 to 29 could be detected.
(2) Site Index ranged from 14 to 26 (based on MTH at age 20).
(3) Basal area/ha ranged from 2.9 to nearly 100 m^2/ha.
(4) Stocking ranged from 200 to 2625 stems per hectare with the majority between 400 to 1000. There was a gap for younger stands with high stockings.

3.4.2 Data for model validation

There were 505 plot measurements within 236 plots reserved for model validation. A summary of plot variable estimates with a breakdown for plains and foothills is listed in Table 3.4. Figure 3.4 displays the growth pattern of main modelling variables in the data set.

Table 3.3: Summary of data for model building

Variable	Plains					Hills				
	Freq.	Mean	Std.D.	Min.	Max.	Freq.	Mean	Std.D.	Min.	Max.
AGE	2305	13.70	4.95	7.45	29	186	15.47	4.56	7.8	25.7
MTH	2305	13.97	4.24	5.79	26.66	186	16.74	5.23	7.43	31.53
G	2305	24.95	13.90	2.90	74.12	186	45.18	21.58	8.41	98.91
N	2305	722	213	200	2625	186	952	465	375	2075
V	2305	145.4	112.8	8.0	575.0	186	289.6	171.7	26.8	722.1
Altitude	2305	130	61	40	384	186	426	74	300	550
Rainfall	2305	726	67	635	990	186	983	55	900	1020
Site index	2305	20	2	14.5	23	186	22	2	18	26
Pruning height	2305	2.48	0.58	0	4.38	186	1.99	1.66	0	6
GF rating	2305	5.3	3.5	2	9	186	3.5	2.9	2	9

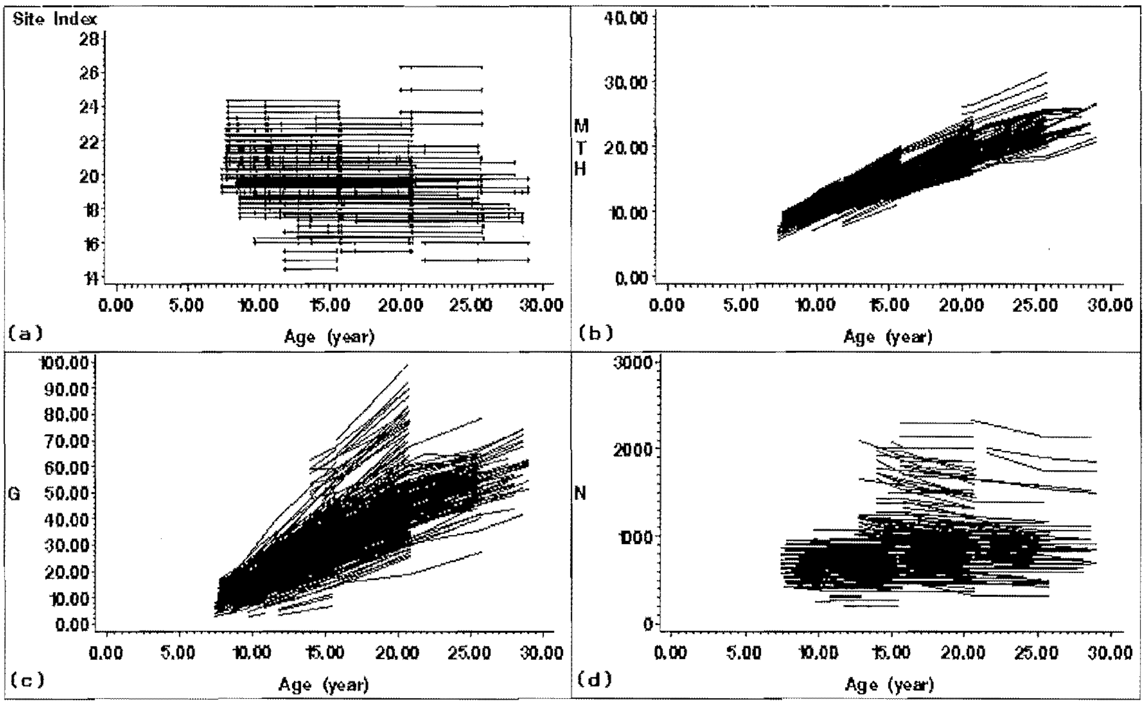


Figure 3.3: Growth patterns of main modelling variables in data set for model building

(a) site index, (b) mean top height, (c) basal area, and (d) stems per hectare

Table 3.4: Summary of data for model examination

Variable	Plains					Hills				
	Freq.	Mean	Std.D.	Min.	Max.	Freq.	Mean	Std.D.	Min.	Max.
AGE	495	16.21	4.52	8.45	28	10	15.05	3.71	10.45	20.45
MTH	495	16.71	3.84	6.83	25.56	10	16.86	3.53	12.14	21.65
G	495	32.03	12.71	5.00	65.61	10	42.92	17.99	19.43	73.62
N	495	699	162	225	1175	10	660	125	525	850
V	495	205.2	112.9	15.3	535.6	10	271.8	145.7	91.3	523.1
Altitude	495	119	56	47	235	10	430	77	370	520
Rainfall	495	713	60	641	886	10	1020	0	1020	1020
Site index	495	20	2	16	24	10	22	1.22	21	24
Pruning height	495	2.57	0.37	0.00	4.38	10	2.54	0.05	2.5	2.6
GF rating	495	4.4	3.3	2	9	10	6.20	3.61	2	9

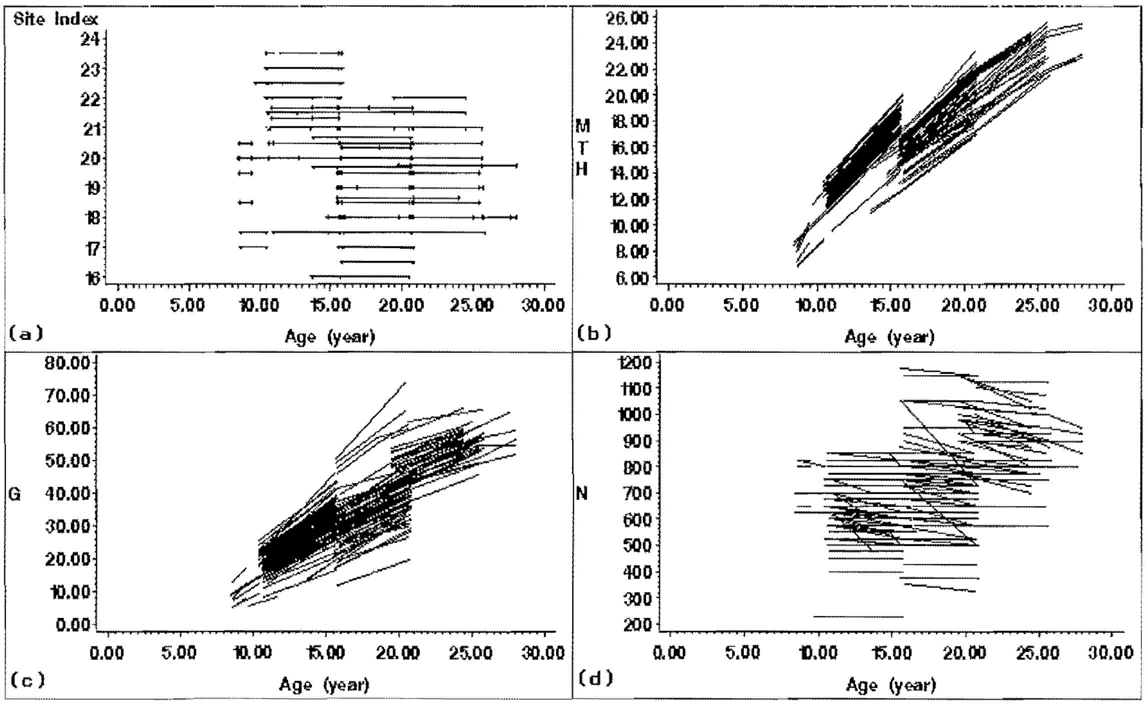


Figure 3.4: Growth pattern of main modelling variables in data set for model validation

3.4.3 Data addition to the database for model building

A few additional data were added to the existing database due to a lack of enough data from young stands (before age 7) and from old stands (greater than 29 years). The new juvenile growth measurements were data from 4 experiments and they were located in Dalethope, Smith's block and Eyrewell Forest (one established by FRI, and the others by Mason). The older measurements were located in Eyrewell Forest, now owned by Carter Holt Harvey Limited (CHH), which was chosen because it is geographically close to plantations of SPBL. A summary of all the data available for building models is listed in Table 3.5 and the resulting graphical pattern is shown in Figure 3.5.

It was believed that building a whole rotation model by joining the juvenile growth data with older growth data was more meaningful than linking separately established models of juvenile growth and older. Measurements for juvenile growth were taken between ages 3 and 6, by which time the majority of trees have reached breast height (1.40 m). They were taken as passive data and the data with significant treatments were dropped. The inclusion of juvenile measurements would improve model projection ability for young trees that were not older than age six.

Juvenile trees were all measured before thinning and pruning, but the rest of the data were all after thinning and pruning. For cross projections from an initial age (T_1) younger than age six to a future age (T_2) older than age six, that thinning would not significantly change the growth rate should be assumed under the following conditions:

- No thinning from above occurred so mean top height might not change. It was usual in SPBL to practise thinning from below.
- The remaining stocking was high and pruning height was low enough so that no significant changes in environment and no severe crown loss were caused. In SPBL the remaining stocking after the first thinning was higher than 500 and pruning height was mostly 2.5 m.
- Competition between trees was not serious at the stage of first thinning around age seven to nine years. The wider gaps created by thinning might not significantly affect growth rate compared with narrower gaps if planting stocking was not high.

These assumptions might not be strictly followed for many stands. It was expected that a better model could be developed by including these additional data than excluding them.

Table 3.5: Summary of data with additional measurements for model building

Variable	SPBL			Juvenile			Eyrewell		
	Freq.	Mean	Range	Freq	Mean	Range	Freq.	Mean	Range
AGE	2491	13.8	7.5~29	44	4.2	3~6	129	23.1	14.3~31.4
MTH	2491	14.2	5.8~31.5	44	3.6	1.9~5.1	129	21.7	15.1~28.1
G/ha	2491	26.5	2.9~98.9	44	2.2	0.06~5.7	129	32.8	9.9~67.1
N/ha	2491	739	200~2625	44	1390	615~2500	129	594	200~1831
V/ha	2491	156.1	8~722	-	-	-	-	-	-
Altitude	2491	151.8	40~550	44	249	80~500	129	149	76~198

Freq. is the number of plot measurements, "-" missing values

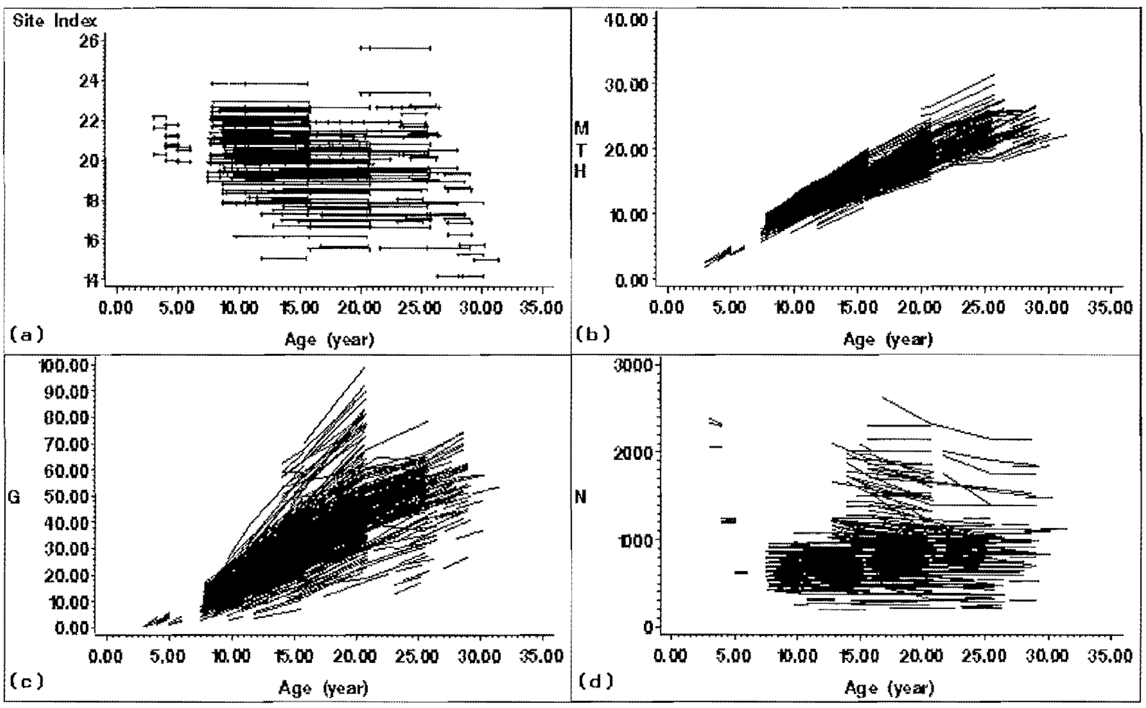


Figure 3.5: Growth patterns of main modelling variables with additional data

The newly added data could be used for fitting equations to MTH, basal area/ha and stems/ha but not for other components. The trees of new juvenile re-measurements were not clearly identified in later records and the suitability of the existing tree volume equation was not established for the very young trees. Only plot statistics of MTH, basal area/ha and stems/ha were provided by CHH. Data used for each component therefore were slightly different due to variations in data availability and relevance.

Individual-tree information is not included in this section, but is described in later chapters.

Altitude, rainfall and geographical type (plains and foothills) were searched as the main environmental variables. Altitude was measured with a map of SPBL. Annual rainfall observations were obtained from the National Institute of Water and Atmospheric Research Institute (NIWA). Rainfall showed a linear relationship with altitude, which is seen in Figure 3.6. Genetic Growth and Form rating (GF) was at only two levels: a value of two was for stands planted before 1979 and 9 for after 1979. Pruning height was from 0 to 6 meters.

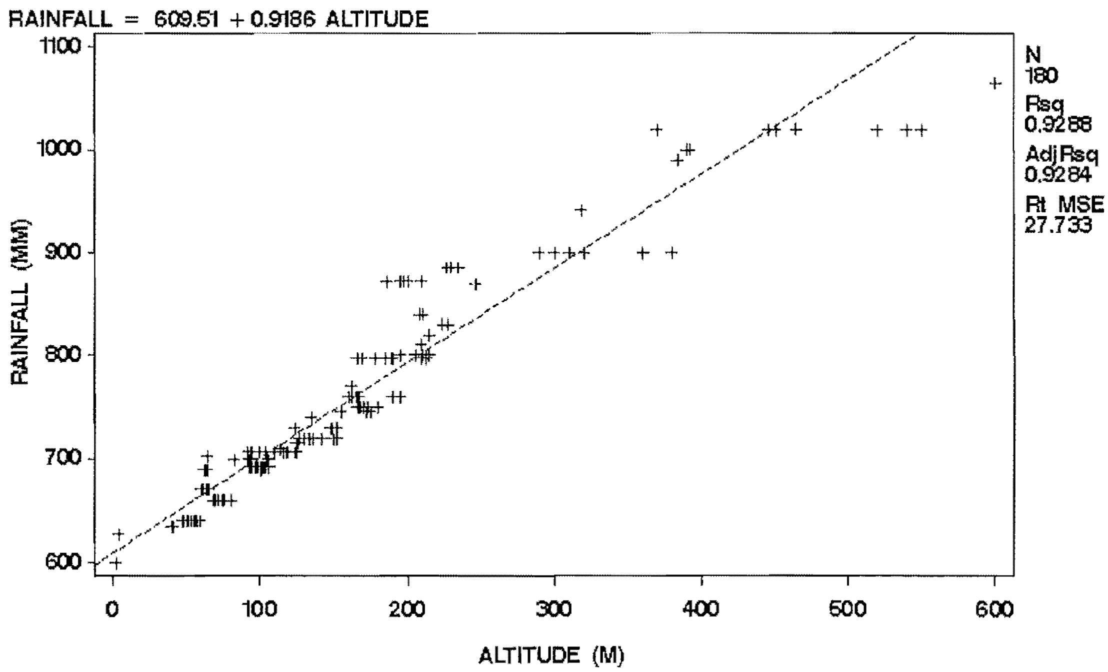


Figure 3.6: The relationship between altitude and average rainfall

3.5 SUMMARY

A new database for permanent sample plot measurements was established using a database management system. Multiple levels of tables and attributes in each table were identified. A relational database was structured and used to manipulate data. Main variables required for growth modelling were computed using data from the database.

The main variables for growth modelling were described in both tabular and graphical forms. The main data set comprised a total of 1200 plots with 2664 repeated measurements for modelling. Stand age ranged from 3 to 31 years, site index from 14 to 26 m at age 20, basal area/ha from 0 to a maximum of 100 m^2/ha, and stocking/ha from 200 to 3000 stems/ha with the majority between 400 to 1000 stems/ha. The database was partitioned for model building and model validation. Extra juvenile growth data were added to the data set for model building.

Starting with the establishment of a database enabled analysis for this study to avoid potential misunderstandings when using data provided by a third party. The modeller's own database was organised specifically according to research needs. The next chapter describes the modelling of height-diameter relationships, which was needed to estimate individual-tree height and mean top height for stands.

CHAPTER 4

MODELLING OF HEIGHT-DIAMETER RELATIONSHIPS

This chapter describes the modelling of height-diameter relationships of *Pinus radiata* at stand and regional levels in Canterbury.

4.1 INTRODUCTION

Height-diameter relationships play an important role in forest mensuration systems. Stand mean height, mean top height, and mean height for each diameter class are usually calculated directly from equations of tree height (h) on diameter at breast-height outside of bark (d). Volume per hectare is an accumulation of individual tree volumes often derived from a two-dimensional function using measured d and estimated h.

Height-diameter equations, which pass through the origin (d=0 cm, h=1.40 m), attain an asymptote for large values of diameter and always have a positive slope, have been recognised by a number of authors as the most appropriate form (for example, Curtis, 1967; Garman *et al.*, 1995). The following Petterson equation (Schmidt, 1967) is an example of a height-diameter equation with these properties. It can be seen that h is 1.40 m when d approaches zero; the asymptote is $1.40+\alpha^{-2.5}$ when d approaches $+\infty$; and the derivative of the curve is $2.5\beta/[d^2(\alpha+\beta/d)^{-3.5}]$, that is always larger than zero since d, α and β are all positive in the equation.

The Petterson equation with exponent –2.5: $h=1.40+(\alpha+\frac{\beta}{d})^{(-2.5)}$

where h is tree total height in m, d is diameter at breast height in cm, α and β are parameters.

Linearity is another good property of equations, especially in permanent sample plot (PSP) database computational systems, because coefficients can be solved explicitly and uniquely. Early height-diameter equations were forms of linear and linearized equations (Henriksen, 1950; Myers, 1966; Curtis, 1967), many of them are still in use today. The above Petterson equation can be easily transformed into linear form and has been widely used in New Zealand (McEwen, 1978, Goulding, 1995). The fit it provides for radiata pine on varying sites and ages needed to be examined, however. As the power of computers developed, a number of non-linear equations were tried because they are often more flexible than linear equations (Larsen and Hann, 1987; Wang and Hann, 1988; Arabatzis and Burkhart, 1992; Huang *et al.*, 1992; Dolph *et al.*, 1995).

A height-diameter equation may be created and applied at a plot level (Curtis, 1967; Garcia, 1974), a stand level in which a few plots are sampled, or a local region (Larsen & Hann 1987; Wang & Hann 1988, Huang *et al.*, 1992). It is common practice in some forestry companies, Selwyn Plantation Board Limited (SPBL) for example, that each height-diameter equation be determined at the stand level rather than just for a single plot. For radiata pine plantations, a stand is a continuous area of forest which is distinguished from its surroundings but relatively uniform internally in stand age, genetics, site conditions and management regimes. Robust equation forms which fit well under a wide range of stands are desirable.

When a regional height-diameter model is used, no samples of height measurements are required. A better prediction of height might be obtained by identifying and incorporating relevant factors accounting for difference among stands into a model of height on diameter rather than employing complicated functional forms which use diameter only. The relationship between height and diameter in a region may vary with tree species, stand age, site characteristics, genetics, stocking, and silvicultural treatment (Trorey, 1932; Buford, 1986; Larsen and Hann, 1987; Wang and Hann, 1988; Dolph, 1989; Knowe, 1994).

The objective of the study reported here was firstly to determine for various radiata pine stands the most suitable forms of height-diameter equations in terms of the desirable properties and fitting performance, and secondly to build a regional height prediction model for radiata pine growing in Canterbury, New Zealand.

4.2 DATA

The data used here were all from permanent sample plots within SPBL's estate and they involved all 529 repeated stand measurements within 168 stands. The number of plots per stand ranged from 1 to 41. Re-measurement intervals were 3 years and repeated 3 times, on average. About 95% of plots were 0.04 ha in area, while the remainder were 0.02 ha. Table 4.1 summarises the main variables.

To evaluate and test equations, data were prepared in three sets. The first set contained all 529 repeated measurements for 168 stands, in which the number of height-diameter records varied between 7 and 328 trees (i. e. from 1 to 41 plots). This first set was used for fitting height-diameter equations at a stand level. The second set comprised 168 independent stand measurements; one and only one measurement had to be chosen randomly from several measurements in each stand. This data set had no autocorrelation and was used for hypothesis testing of regression coefficients. The third data set was the full data set totalling 24 790 trees without any grouping by stand or by date of measurement, used for building the regional model.

Table 4.1: Summary of data for modelling height-diameter relationships

Variable	Mean	Std. Dev.	Min.	Max.	Frequency
Tree dbh (cm)	21.5	7.0	1.5	58.7	24790 *
Tree height (m)	13.7	4.7	1.7	36.4	24790 *
Stand age (year)	14.5	5.26	7.45	29	529 **
Site Index	20	2	15	28	529 **
Stocking (stems/ha)	752	246	300	2225	529 **
Altitude (m)	148.5	92.7	2	550	168 **
Numbers of plots per stand	6	6	1	41	168 **
Numbers of repeated measurements per stand	3	0.612	2	5	168 **

* Pooled statistics representing regional level

** Statistics at stand levels

4.3 EQUATIONS

Sixteen height-diameter equation forms (listed in Table 4.2) were identified as having potential and then evaluated. Most have desirable properties regarding the origin, slope and asymptote, as outlined above. All equations have positive slopes and all but Equations 4.11 and 4.12 in Table 4.2, recommended by Garcia (1974) for radiata pine in Chile, pass through the origin. All except Equations 4.9 and 4.10 have an asymptote at large values of diameter. None of them is a linear equation but many of them, the Petterson equation for example, can be transformed to a linear model. Equations 4.13 to 4.16 are three-parameter forms, and all others contain only two parameters.

Table 4.2: Height-diameter equations and example references

No.	Equation form	Example reference
4.1	$h=bh+\alpha\, d/(\beta+d)$	Bates & Watts, 1980; Huang *et al.*, 1992
4.2	$h=bh+d^2/(\alpha+\beta\, d)^2$	Huang *et al.*, 1992
4.3	$h=bh+(\alpha+\beta/d)^{-2.5}$	Schmidt, 1967 (exponent n = –2.5)
4.4	$h=bh+(\alpha+\beta/d)^{-5}$	Schmidt, 1967 (exponent n = –5.0)
4.5	$h=bh+(\alpha+\beta/d)^{-8}$	Schmidt, 1967 (exponent n = –8.0)
4.6	$h=bh+\alpha(1-\exp(\beta d))$	Meyer, 1940; Huang *et al.*, 1992
4.7	$h=bh+\alpha(1+1/d)^{(-\beta)}$	Curtis, 1967; Huang *et al.*, 1992
4.8	$h=bh+\exp(\alpha+\beta/d)$	Schumacher, 1939; Curtis, 1967; Buford, 1986; Arabatzis & Burkhart, 1992; Huang *et al.*, 1992
4.9	$h=bh+\alpha(\ln(1+d))^{\beta}$	Modified from an equation in Curtis, 1967
4.10	$h=bh+\alpha d^{\beta}$	Arabatzis & Burkhart, 1992
4.11	$h=\alpha+\beta\exp(-0.08\, d)$	Garcia, 1974
4.12	$h=\alpha+\beta\,(d+10)^{-1}$	Garcia, 1974
4.13	$h=bh+\alpha(1-\exp(\beta\, d))^{\gamma}$	Richards, 1959; Huang *et al.*, 1992; Garman, 1995
4.14	$h=bh+\alpha(1-\exp(\beta d^{\gamma}))$	Yang *et al.*, 1978; Huang *et al.*, 1992
4.15	$h=bh+\alpha/(1+\beta^{-1}\, d^{-\gamma})$	Huang *et al.*, 1992
4.16	$h=bh+\exp(\alpha+\beta\, d^{-\gamma})$	Larsen & Hann, 1987; Flewelling & De Jong, 1994; Lappi, 1997

Note: h = height of trees in m, bh = breast height (1.40 m in NZ); d = diameter at breast height in cm; α, β, γ = parameters in the equations.

4.4 Methods

Each height-diameter equation for each individual stand measurement was fitted with non-linear procedures of the SAS software that employ least-square methods (SAS Institute Inc., 1990). A lower mean square error (MSE) usually means a better fit and less bias for a group of data. MSE might vary from stand to stand, however, due to variations in age, site, or the number of trees per stand. A good equation should fit well all those individual stand measurements. The mean of mean square errors (MMSE) that was calculated as the quotient of sum of the MSE for all stand measurements to the number of total stand measurements ($\sum$ MSE/529) was a logical measure of overall performance of the equation. MMSEs for all equations were compared and ranked.

A regional model was derived by identifying and incorporating significant site factors and stand variables into a model based on the stand-level findings. Variables other than tree diameter might affect height predictions substantially. The independent data set (free of autocorrelation) was used in regression analyses of estimated coefficients to test for the significance of available variables such as stand age, site index, stocking and altitude. Data with autocorrelation due to repeated measurements may produce unbiased coefficients using classical least-squares regression, but the variances of error terms and variances of coefficients may be underestimated. Hypothesis tests based on autocorrelated data are therefore invalid (Neter and Wasserman, 1974; West *et al.*, 1984; West, 1995). Finally a non-linear procedure with the full data set was employed to refine parameter estimates for the tested model (Mason, 1992; Zhao and Mason, 1996). Graphical patterns of residuals were used to detect bias.

4.5 RESULTS AND DISCUSSION

4.5.1 Model fitting at a stand level

The means of mean square error (MMSE) for each equation are listed in the third column in Table 4.3 and they are ranked in the fourth column.

Among two-parameter equations (Equations 4.1 to 4.12), Equations 4.9 and 4.10 led to the largest MMSE probably because they lack an asymptotic property. The rest of the MMSEs were quite close but Equations 4.3 to 4.8 resulted in better fits for all stands.

I. Equation 4.6 gave the least MMSE (1.3670) among two-parameter equations. It is a form of the two-parameter Richards or Weibull equation (when the third parameter equals one). It has the properties of passing through the origin and approaching an asymptote but it is a non-linear form and cannot be transformed to linear form.

Table 4.3: Equation fitting results for h-d relationships

No.	Equation form	Stand level		Regional level	
		MMSE	Rank	MSE	Rank
4.1	$h=bh+\alpha\, d/(\beta+d)$	1.3750	10	6.0708	9
4.2	$h=bh+d^2/(\alpha+\beta\, d)^2$	1.3695	8	6.0310	6
4.3	$h=bh+(\alpha+\beta/d)^{-2.5}$	1.3689	5	6.0553	7
4.4	$h=bh+(\alpha+\beta/d)^{-5}$	1.3684	2	6.1394	11
4.5	$h=bh+(\alpha+\beta/d)^{-8}$	1.3685	3	6.1837	12
4.6	$h=bh+\alpha(1-\exp(\beta d))$	1.3670	1	6.0702	8
4.7	$h=bh+\alpha(1+1/d)^{(-\beta)}$	1.3686	4	6.2390	13
4.8	$h=bh+\exp(\alpha+\beta/d)$	1.3689	5	6.2740	14
4.9	$h=bh+\alpha(\ln(1+d))^{\beta}$	1.3870	11	6.0277	5
4.10	$h=bh+\alpha d^{\beta}$	1.4024	12	6.0874	10
4.11	$h=\alpha+\beta\exp(-0.08\, d)$	1.3694	7	7.1743	16
4.12	$h=\alpha+\beta\,(d+10)^{-1}$	1.3704	9	6.8519	15
4.13	$h=bh+\alpha(1-\exp(\beta\, d))^{\gamma}$	-	-	6.0107	2
4.14	$h=bh+\alpha(1-\exp(\beta d^{\gamma}))$	-	-	6.0081	1
4.15	$h=bh+\alpha/(1+\beta^{-1}\, d^{-\gamma})$	-	-	6.0116	3
4.16	$h=bh+\exp(\alpha+\beta\, d^{-\gamma})$	-	-	6.0241	4

- At the stand level, too few data were available in several stands to allow convergence in three-parameter equations 13 to 16. It resulted in incompatible comparison.

II. Equations 4.1 and 4.2 can be transformed to the same form as Equations 4.3 to 4.5 but with different exponents. In other words, Equations 4.1 to 4.5 have the same form with exponents -1, -2, -2.5, -5, -8, respectively, and they were all collectively named the Petterson equation family here to distinguish them from other forms.

Equation 4.4 with exponent -5 led to the second smallest MMSE (1.3684) among all two-parameter equations. The Petterson equation family has the desirable properties of origin and asymptote and its equations can be transformed to a linear form, which is very useful in PSP databases because immediate and unique solutions can be obtained with simple algorithms. Equation 4.4, for example, can be rewritten in two ways either as Equation 4.17 or Equation 4.18. Both of them can be transformed into linear models. Equation 4.19 is derived by writing the left-hand side of Equation 4.18 as new variable Y.

$$\left(\frac{1}{h-1.40}\right)^{0.2} = \alpha + \frac{\beta}{d} \tag{4.17}$$

$$\frac{d}{(h-1.40)^{0.2}} = \alpha\ d + \beta \tag{4.18}$$

$$Y = \alpha\ d + \beta, \qquad \text{where } Y = \frac{d}{(h-1.40)^{0.2}} \tag{4.19}$$

III. Equation 4.7 fitted the data equally as well as Equation 4.8, a form of Schumacher's equation (Schumacher, 1939). Both equations have the properties of passing through the origin, approaching an asymptote, positive slope, and being transformable to a linear pattern. Ek (1973) indicated in a study of height-diameter relationships with sampling simulation for small samples that Equation 4.8 performed the best.

IV. Garcia's Equations 4.11 and 4.12 fitted nearly as well as Equations 4.3 to 4.8 but are not recommended for general use. There are potential biases with extrapolation for small trees, as these equations cannot pass through the origin.

For three-parameter equations at the stand level, a convergence problem with the non-linear least-squares procedures was encountered; thus no stable solution could be found for some stand measurements. The lack of convergence for several stands resulted in MMSE estimates that could not be legitimately compared with those of other equations.

In summary, though the fitting differences were small, Equations 4.9 and 4.10 fitted the data poor and Equations 4.6 and 4.4 were the two equations that led to the smallest

MMSE. In particular, Equation 4.4 was the best model that can be transformed to linear form.

4.5.2 Model fitting at the regional level

Equation 4.4 was chosen in the study as a base for the regional model and for creating parameter estimates for all individual stands because it had the smallest MMSE of those equations that could be transformed to a linear form. Linear regression can work more efficiently and produce parameter estimates that usually have smaller variation and with residuals that are closer to normal distributions than can non-linear regression. Equation 4.19, a linear transformation of Equation 4.4, was used to create parameter estimates for every individual stand. The estimated values of parameters were assumed to vary with stand age, site index, stocking (stems per hectare), and altitude. A parameter-predicting model with a linear combination (4.20) was constructed and hypothesis tests were conducted with ordinary linear regression.

$$\left.\begin{aligned} \alpha &= a_0 + a_1 T^{-0.5} + a_2 Alt + a_3 \ln(SI) + a_4 N + a_5 T \times N \\ \beta &= b_0 + b_1 T^{-0.5} + b_2 Alt + b_3 \ln(SI) + b_4 N + b_5 T \times N \end{aligned}\right\} \qquad (4.20)$$

where $a_1 \ldots a_5$ and $b_1 \ldots b_5$ are coefficients in the model, T is stand age, Alt is altitude, N is stocking per hectare, SI is site index, α and β are estimated parameters in the base Equation 4.19.

The outputs of separate regression tests for both parameter α and β with data set two (that was free of autocorrelation) are shown in Table 4.4. Age and site index were significantly related to the α parameter estimates, while altitude was significantly correlated with the second parameter β. The asymptotic parameter α increases with increasing stand age and site index. Stocking (representing stand density) might be expected to affect the relationship between height and diameter but it was not found to be statistically significant.

Table 4.4: Regression testing results for parameter estimates vs. explanatory variables

		Parameter α (R^2=0.83)				Parameter β (R^2=0.10)			
Variable	DF	Coefficient	Std. Error	T	Prob>|T|	Coefficient	Std. Error	T	Prob>|T|
Intercept	1	0.80979	0.04958	16.34	0.0001	0.26516	1.09710	0.242	0.8093
Age	1	0.58063	0.11016	5.271	0.0001	-0.90022	2.43785	-0.369	0.7124
Altitude	1	0.00001	0.00001	1.312	0.1915	0.00086	0.00025	3.406	0.0008
Site Index	1	-0.13844	0.01257	-11.01	0.0001	0.34307	0.27826	1.233	0.2194
Stocking(N)	1	-0.00004	0.00004	-1.101	0.2726	-0.00028	0.00082	-0.345	0.7302
Age × N	1	0.00017	0.00015	1.118	0.2652	0.00053	0.00338	0.157	0.8756
Error	161								

The final step in developing the regional model was to refit the chosen model by including stand age, site index and altitude with all available data (data set three). Two approaches to obtaining the parameter estimates were tried, linear and non-linear regression. Non-linear procedures can often lead to a smaller MSE and avoid back-transforming bias in prediction. Non-linear procedures in the study resulted in Equation 4.21. When dbh alone was used to estimate h (before age, site index and altitude were entered), mean square error was 6.1394. A huge reduction of 50.9% of MSE was gained when age was introduced. A considerable further reduction of 39.5% of MSE was gained through the contribution of site index. Altitude was less important but statistically significant, further reducing MSE by 6.4%. The mean square error with Equation 4.21 was 1.7052, which was 28% of the MSE before age, site index and altitude were introduced. Mean of residuals was -0.0076 and skewness was -0.1533. Figure 4.1 shows the pattern of residuals against predicted values, age, altitude and site index. There was little apparent bias. Ninety-eight percent (>1% and <99%) of residuals were within ±3.5 m and 90% (>5% and <95%) were within ±2.0 m.

$$h = 1.40 + [0.695955 + 0.666983\, T^{-0.5} - 0.106771\, \ln(SI) + (0.954201 + 0.000741 Alt)/d]^{-5} \qquad (4.21)$$

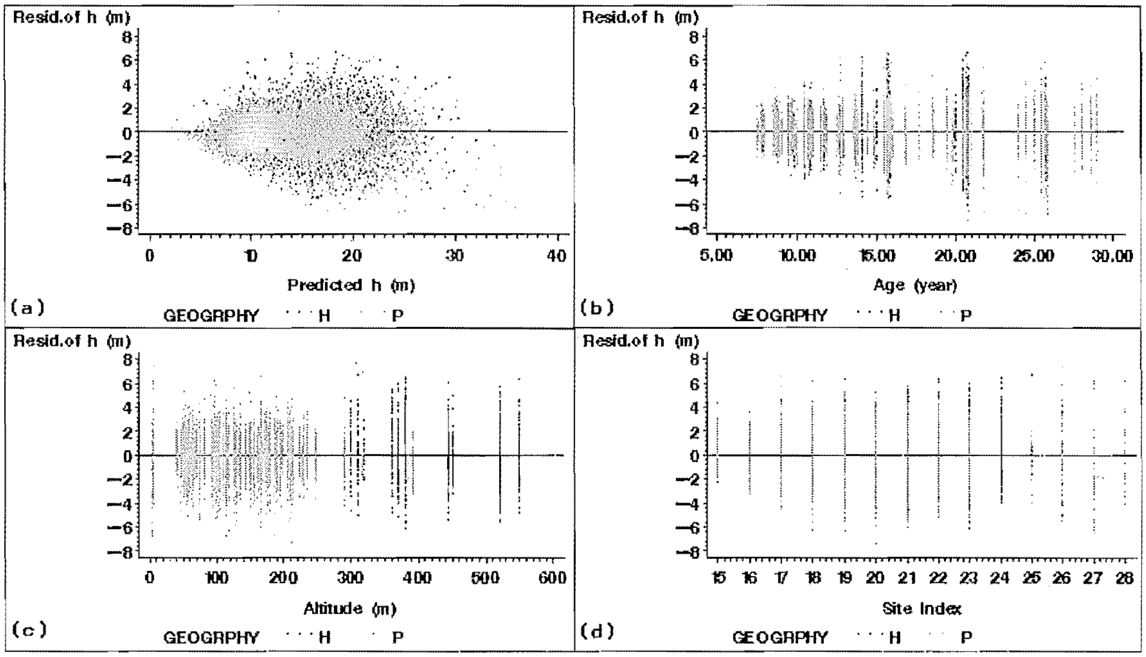


Figure 4.1: Residual pattern of h-d model at a regional scale

With this regional model, tree height can be estimated with tree diameter, stand age, site index and altitude. No samples of height measurements are needed when using the model. The model reflected relationships of radiata pine within SPBL's estate. Its suitability for trees outside the region needs to be investigated. All trees included in the study were older than seven years. Model performance for estimating mean top height in particular could have been examined if actual mean top height was known.

All height-diameter equations estimating height from diameter alone were fitted with the pooled regional data set (Data set 3) and the fitted mean square errors and rankings are listed in the last two columns of Table 4.3. Three-parameter equations resulted in smaller MSE but the potential to reduce MSE was rather limited. These results are in agreement with those of Huang *et al.* (1992) for a regional study of height-diameter relationships in which the MSEs of three-parameter equations were not much different from those of the best two-parameter equations. Variables other than tree diameter should be sought to improve predictions.

4.6 CONCLUSIONS

Of the equations evaluated, fitting differences at stand level were small but the two-parameter Richards' equation and the Petterson equation with exponent -5 led to the smallest means of mean square error for stands varying considerably in age, site index, altitude and the number of trees sampled. The latter equation form, in particular, was the best with the desired properties of passing through the origin, approaching an asymptote, having a positive slope, being transformable to a linear form and leading to small fitted mean square error. Three parameter equations offered limited benefits for many stands at huge computational expense and failed to converge when fitting for a few stands.

A height model at a regional scale was obtained by identifying and incorporating into the selected Petterson equation the most important variable, stand age, the second most important variable, site index, and a less important but statistically significant variable of altitude. The inclusion of these three variables resulted in a reduction of mean square error of 72%. It was expected that 90% of predictions should be within ±2.0 m and the model could be useful especially when no height measurements are available.

CHAPTER 5

EVALUATION OF SAMPLING IN SPBL'S INVENTORY SYSTEM

5.1 INTRODUCTION

Forest inventory is used to provide information about the current resource and changes in it over time. It is usually impossible to measure all trees in a whole forest or stand so sampling techniques are used to provide robust estimation efficiently (Freese, 1962; Husch, 1971; Husch *et al.*, 1982; Goulding and Lawrence, 1992; Shiver and Borders, 1996). The most important principle of sampling is to achieve both estimation reliability and cost-effectiveness with appropriate sampling designs and intensity. Estimation reliability can be measured by means of sampling precision.

For the last 30 years forest inventory has been practised by SPBL primarily for timber production management. The system combines continuous forest inventory (CFI) at frequent intervals with intensive sampling of full PSPs for estimating stand level statistics. The inventory system was established by SPBL during the 1960's and re-measurements were taken regularly (every three years, on average). Permanent sample plots were fully incorporated and no temporary plots were included. Stands rather than the whole forest were regarded as basic populations for sampling. Intensive management system based on stands requires reliable information for each stand.

Sampling precision achieved over the last 30 years of inventory practice in SPBL was of interest to the company. An examination of sample size determination would help SPBL managers in planning future inventories.

The objectives of the evaluation of inventory sampling described here were to further understand the data quality for growth modelling and to reveal sampling precision gained in the past by SPBL. Design of an appropriate inventory system was beyond the scope of the study.

5.2 DATA FEATURES

The general features of the database have been described in Chapter 3. This analytical review of SPBL's sampling refers only to this organisation's plantations.

A simple random sampling with plots bounded and plot-area fixed was used for all stands in SPBL. Sampling intensity was about one per cent and sample size (the number of plots in a stand) varied from 1 to 41. Altogether, there were 513 repeated stand measurements in 165 stands, as listed in Table 5.1. All measurements for which only one plot was sampled within a stand were not included in the evaluation procedure, as variance among plots could not be estimated.

Table 5.1: Actual sample size frequencies in SPBL

Sample Size	2	3	4	5	6	7	8	9	10	15	20	25	30	35	40+	Total
Frequency	80	108	95	60	43	27	11	14	27	22	12	5	3	2	4	513
Percent(%)	16	21	19	12	8.4	5.3	2.1	2.7	5.2	4.3	2.3	1	0.6	0.4	0.8	100

Two trends were clearly revealed relating to determination of sample size in the system.

Firstly, sample size was proportional to stand area (see Figure 5.1) which was directly caused by the practice of a fixed percentage sampling intensity. The sampling strategy may have assumed that the coefficient of variation (CV) could be linearly related to stand area.

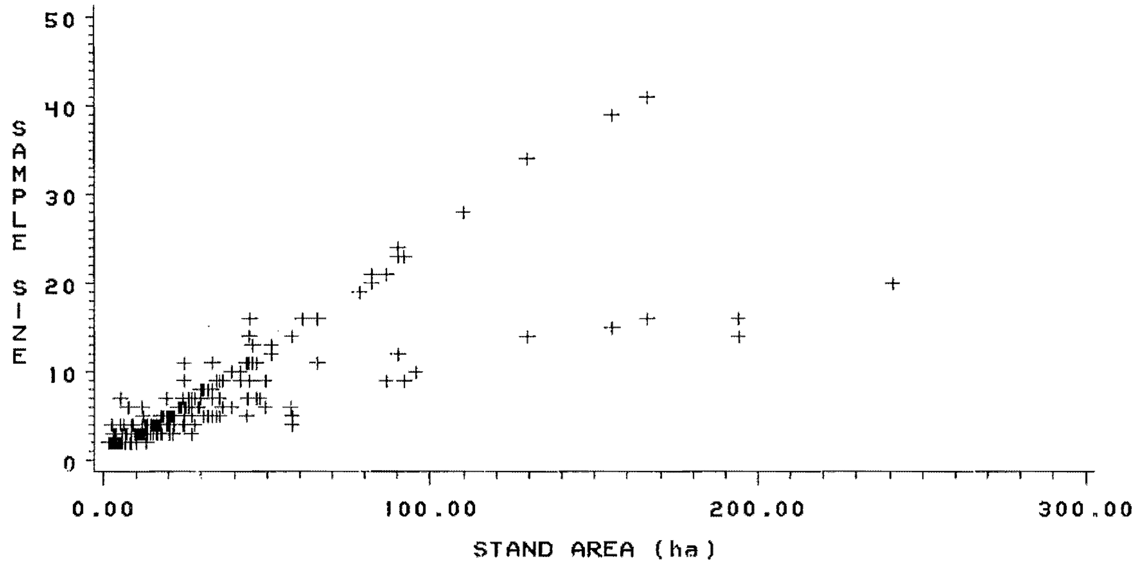


Figure 5.1: Common practice of sample size increase with stand area in SPBL

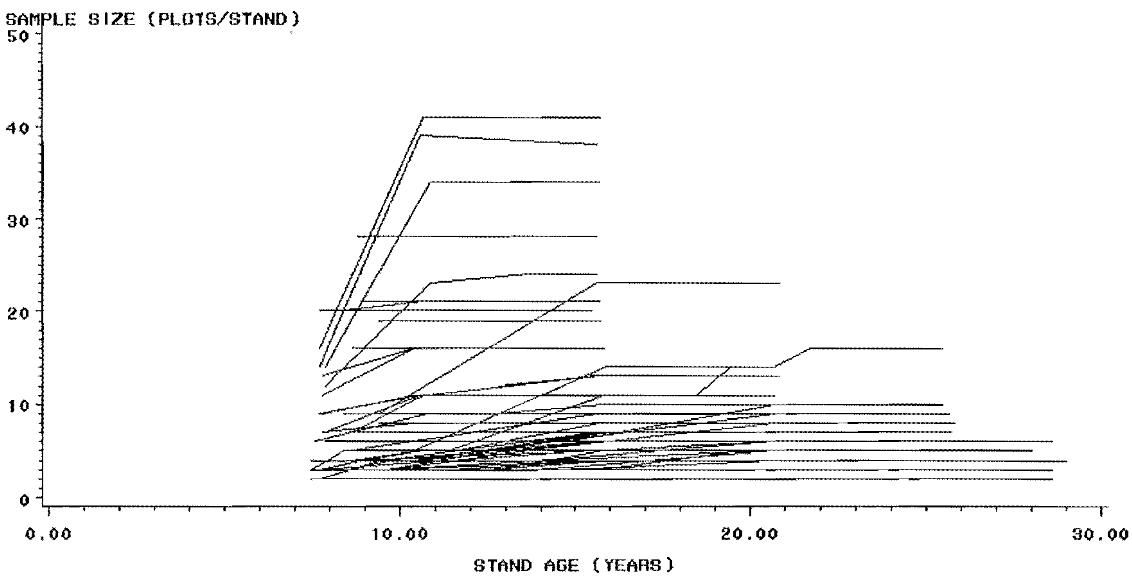


Figure 5.2: Actual sample size changes with age in SPBL's inventory

The second trend was that sample size increased from time to time with stand age (Figure 5.2) possibly under an assumption that CV might increase with stand age. Apparently, CV was not the consideration in determining sample sizes in SPBL's inventory system.

5.3 METHODS

Precision is a useful criterion for assessing sampling reliability. Accuracy is an ideal measure of sampling quality but not all trees were measured in any stand so true values of stand parameters cannot be known. An estimate for a population is expressed as sample mean or range with lower and upper bound of confidence interval, which is described by Equation 5.1. The half 95 percent confidence interval is estimated with Equation 5.2 for sampling without replacement. Error bounds of standard error as a percent of mean can be calculated with Equation 5.3. A measure of sampling precision is then calculated by Equation 5.4. Precision is determined by the ratio of the standard error to the mean, Student's t distribution and sampling fraction (i.e. sampling intensity). A value of positive 100% is an ideal precision. Extremely poor sampling can lead to a negative value that is caused when variation among samples is so high that upper bound of error is larger than the sample mean.

$$\hat{\bar{Y}} = \bar{y} \pm E \tag{5.1}$$

$$E = t_{0.05} s_{\bar{y}} \sqrt{1 - \tfrac{n}{N}} \tag{5.2}$$

$$E\% = \frac{E}{\bar{y}} 100 \tag{5.3}$$

$$P\% = (1 - \frac{E}{\bar{y}})100 = (1 - \frac{t_{0.05} s_{\bar{y}}}{\bar{y}} \sqrt{1 - \tfrac{n}{N}})100 \tag{5.4}$$

where $\bar{y}$ = estimate of mean value of population

$s_{\bar{y}}$ = standard error of $\bar{y}$

n = actual sample size

N = number of sample units in a population, the ratio of stand area to plot area

$t_{0.05}$ = t value of Student distribution with confidence 95% and df = n-1

E = upper bound of standard error with percent of confidence

E% = upper bound of standard error as a percent of mean

P% = sampling precision

Sample size can be determined before formal sampling. It is calculated with Equation 5.5 for random sampling without replacement in a finite population. Equation 5.6 is for an infinite population or for sampling with replacement in a finite population or sampling without replacement if sampling intensity is less than 5%. Sample size is chosen in consideration of CV, stand area and plot area, AE (allowable error in Shiver and Borders 1996; or desirable probable limits of error in Goulding & Lawrence 1992). AE is another expression of E% but is subjectively decided by managers, say 5% to 20%, with respect to expected sampling precision 95% to 80%. On the one hand, intensive sampling is required to achieve high precision, and minimum sample size is desirable to reduce expenditure on the other.

$$n = \frac{Sarea(t_{0.05}^2)CV^2}{Sarea(AE)^2 + t_{0.05}^2 CV^2 Plotsize} = \frac{N(t_{0.05}^2)CV^2}{N(AE)^2 + t_{0.05}^2 CV^2} \tag{5.5}$$

$$n = \frac{t_{0.05}^2 (CV)^2}{(AE)^2} = \frac{t_{0.05}^2 s_y^2}{(AE)^2 \bar{y}^2} \tag{5.6}$$

where n = needed sample size

s_y^2 = estimated variance of y

CV = coefficient of variation of y (the ratio of standard deviation s_y to mean $\bar{y}$)

AE = allowable upper bound of standard error as a percent of mean

Sarea = stand area

Plotsize = plot area

To seek a way to obtain CV of volume, the correlation of CV of volume with CV of basal area, stocking and mean top height was analysed. Regression testing for CV of volume was conducted against stand age and other stand condition variables with randomly chosen autocorrelation-free data. Estimation of CV is usually the main technical problem in practice when determining appropriate sampling size, and pilot sampling is needed before formal sampling. Volume is usually the main concern for forest managers.

5.4 Results

5.4.1 Sampling precision of volume, basal area, stocking and MTH

Sampling precision for each stand measurement was calculated at a 95% confidence level and the results are summarised in Table 5.2. Average precision for volume/ha, basal area/ha, stocking/ha, and mean top height was 65%, 66%, 65% and 96%, respectively. When stand area as weighting was considered, the average precision for volume, basal area, stocking, and mean top height increased to 83%, 83%, 84% and 98%, respectively.

Table 5.2: Sampling precision distribution for main stand variables

Variable	Non-weighted Mean	Weighted Mean	Min.	10% Percentile	90% Percentile	Max.
V (volume)	65%	83%	-333%	22%	94%	97%
G (basal area)	66%	83%	-301%	29%	94%	99.98%
N (stocking)	65%	84%	-490%	29%	94%	100%
MTH	96%	98%	49%	89%	99%	99.9%

It is also seen from Table 5.2 that there were 10% of stands for which sampling precision for volume, basal area and stocking were no more than 22%, 29% and 29%, respectively. On the other hand, there were another 10% of total stands for which sampling precision for volume/ha, basal area/ha and stocking/ha were all no less than 94%. Sampling precision for mean top height was consistently high, probably because of the asymptotic property of mean top height. It is apparent that sampling precision was distributed unevenly.

5.4.2 Variation in coefficient of variation (CV) among stand variables

The correlation coefficients among CV of volume/ha, CV of basal area/ha, mean top height, and stems/ha were calculated and the result is shown in Table 5.3. The results of regression testing of CV of volume/ha against such variables as age, stand area, altitude, Genetic GF rating, pruning height, and thinning are listed in Table 5.4.

Table 5.3: The relationship of coefficient of variation among different variables

Variable	Age 7 to 30				Age 7 to 10			
	CV of V	CV of G	CV of N	CV of H	CV of V	CV of G	CV of N	CV of H
CV of V	1				1			
CV of G	0.992	1			0.996	1		
CV of N	0.385	0.457	1		0.535	0.577	1	
CV of H	0.562	0.496	0.057	1	0.612	0.559	0.079	1

Table 5.4: Regression testing results of CV of volume versus explanatory variables

Variable	DF	Parameter Estimate	Std. Error	T value	Prob>|T|
Intercept	1	0.26157	0.03377	7.745	0.0001
Area	1	-0.00004	0.00017	-0.218	0.8279
Age	1	-0.00577	0.00132	-4.38	0.0001
GF rating	1	-0.00557	0.00190	-2.924	0.0040
Pruning Height	1	-0.01200	0.00712	-1.685	0.0939
Altitude	1	0.00017	0.00006	2.823	0.0054
Error	159				

R^2=0.1647

Table 5.3 shows that CV of volume was highly related to CV of basal area with a correlation coefficient of 0.99, as can also be seen in Figure 5.3. CV of basal area could therefore be used as a good estimate of CV of volume. CV of basal area is relatively easy to obtain with several plot measurements. An even easier measure is CV of stocking. The correlation of CV of stocking to CV of volume, however, was not very high and it decreased as age increased. At the first measurement (age 7 to 10), the correlation coefficient was 0.535 and it reduced to 0.385 when overall measurements were considered.

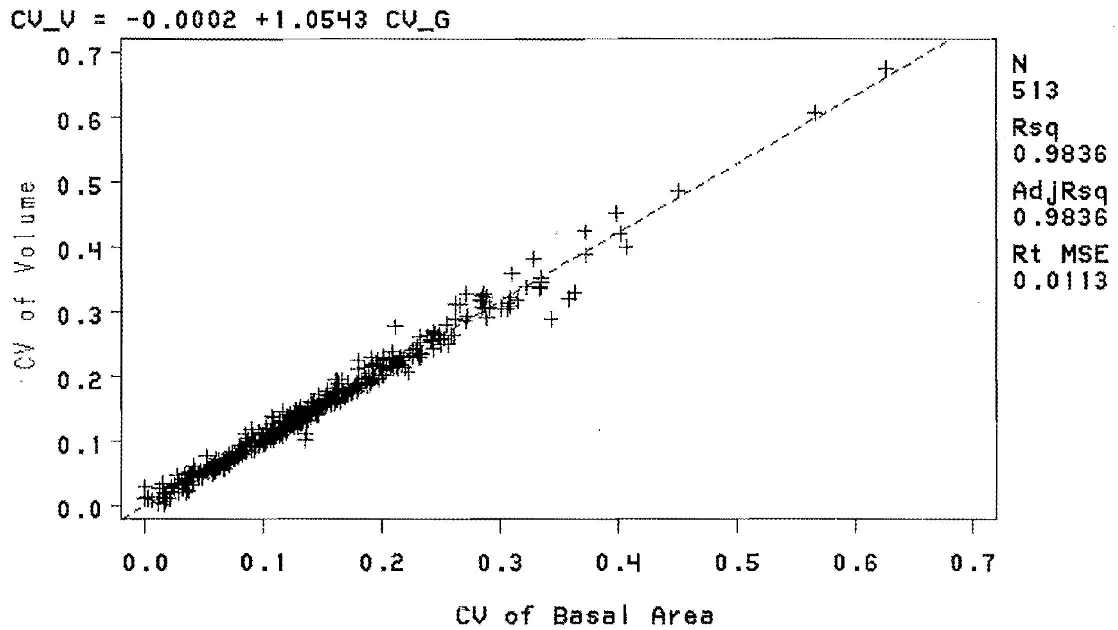


Figure 5.3: Linear relationship between CV of volume and CV of basal area

Regression testing results for a proposed hypothesis that coefficient of variation of volume was correlated to age, stand area, pruning height, Genetic GF rating and altitude are shown in Table 5.4. The results were obtained with a data set in which only one measurement was randomly sampled for each stand to avoid autocorrelation. Age and GF rating were negatively correlated to CV of volume, and altitude was positively correlated to CV of volume. Stand area and pruning height showed no significance at 5% level. CV of volume decreased with age for many stands and increased in only a few cases (Figure 5.4). Site conditions in a stand varied more at higher altitude probably due to the difference in slope, aspect and temperature. Improved GF rating resulted in more uniform tree size in stands. The R-square of this regression was 0.165

only, so CV of volume could not be reliably predicted from these explanatory variables. It was noted that GF rating was a time-dependent factor (classed as value 2 and 9 in terms of planting year before or after 1979); therefore, it might represent not only itself but also such factors related to time as improved seedling handling, cultivation and weeding.

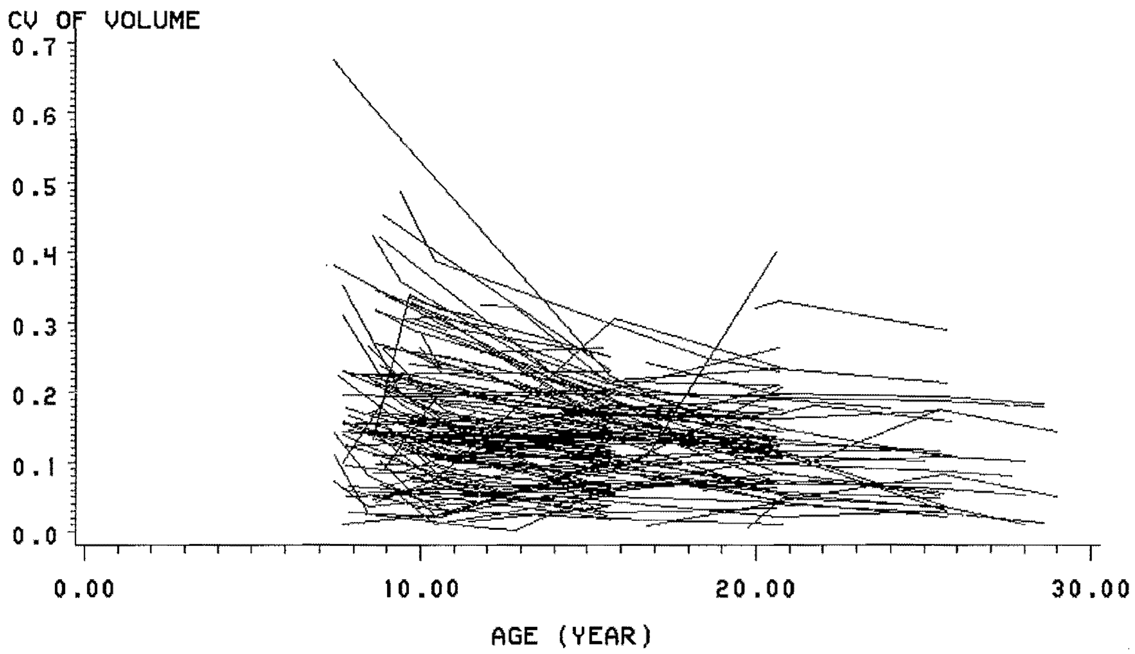


Figure 5.4: CV of volume decreased with age for many stands and increased in few cases

5.4.3 The determination of sample size

Figure 5.5 shows the profile of needed sample size versus actual sample size in SPBL. The required sample size was calculated with Equation 5.5 for all stand measurements, given AE=10% for stand areas greater than 30 ha, otherwise AE=20%. Most efficient sampling occurs when the actual sampling size equals the needed sample size, which follows the reference line. The points over the reference line are those stands which gained a lower sampling precision and the points under the reference line, on the other hand, are the stands with too many plots sampled (i.e. inefficient investments). Without a calculation to determine needed sample size before formal sampling, actual sample size had been more than necessary for some stands and less than necessary for many others.

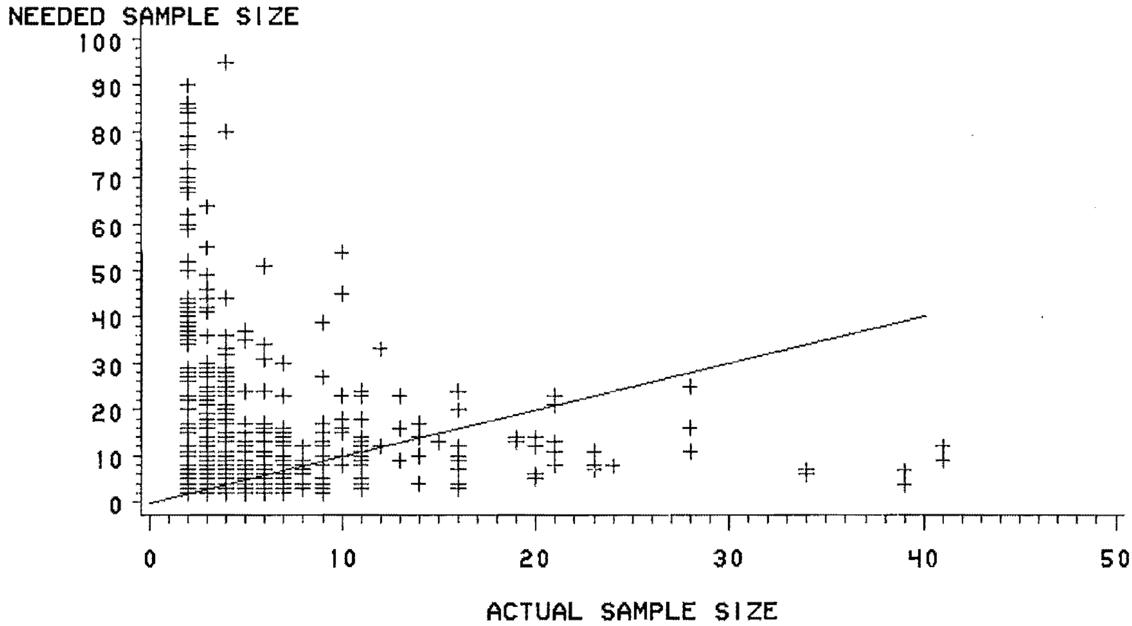


Figure 5.5: The profile of needed sample size versus actual sample size

To determine an appropriate sample size, several considerations should be helpful.

(a) Sampling the first time (at plot establishment) is crucial. To estimate CV, useful approaches can be a pilot survey, a calculation in field after measurements of several formal plots, or use of previous sampling of similar stands.

(b) CV of volume can be reliably estimated by CV of basal area, as both were extremely close (Figure 3.4). Using a point-sampling scheme for the pilot survey at the first measurement should provide an easy solution for obtaining estimates of basal area/ha. Point-sampling is impossible, however, where a dense under-storey causes visual difficulty in using an angle gauge. CV of stems per unit area can also be a good replacement if the tree size is uniform.

(c) AE can be determined in respect to stand area, which could be an explicit way for managers to make decisions. For a large area, say more than 50 ha, probably 5% to 10% of AE could be used to ensure more plots to obtain a good estimate. For a small area, say less than 10 ha, 20% of AE might not be a problem. Adjustment might be needed after the first calculation. For example, if the needed sample size is an excessive number for a very small stand, an artificial rise of AE can reduce the needed sample size enormously. This procedure lowers sample size by reducing sampling precision for the special circumstance. There must be a predetermined limit for AE, however. The cost in measuring a plot should also be considered if it is different among stands.

(d) An adjustment can be made for the second measurement based on the results of sampling the first time. Re-assessment of sample size for a stand is needed following certain artificial or natural disturbances. Stratified sampling may be a more efficient sampling scheme if variations in site and stand density are high and the stand can be divided into several relatively homogeneous strata.

5.5 DISCUSSION AND CONCLUSIONS

Sampling precision for both volume/ha and basal area/ha achieved over the last 30 years of forest inventory practice of SPBL was 65%, on arithmetic average, and 83%, for the average weighted with stand area (calculated at a 95% confidence level). The sampling precision was very unevenly distributed by stands.

This study has revealed major drawbacks to SPBL's CFI system. Inventory practice in SPBL needs to be changed to a system more in accordance with sampling theory. Coefficient of variation and allowable error should be considered when determining required sample size. Coefficient of variation of basal area, which is easier to measure, can provide reliable estimates for coefficient of variation of volume, which is the main statistic of interest to managers. Stand area and cost per sampling unit should be taken into account in determining an appropriate allowable error. Further research into the most efficient inventory design for SPBL is needed and it may include choice of sampling scheme, optimum plot area, appropriate numbers of temporary versus permanent plots, and desirable precision required at different crop ages.

Data requirements for growth modelling demand that individual trees are unambiguously identified (when a tree level model is built), plots are homogeneous, and extremes of site and stand conditions are sampled (Vanclay, 1994). The two former requirements were satisfied in SPBL's inventory system. The third may or may not have been satisfied in the system as a random principle was usually followed in the sampling system. For the sake of growth modelling it is recommended that some plots at the extremes of age, altitude and stocking be subjectively selected in SPBL's estate so that model predictions may be interpolations rather than extrapolations.

CHAPTER 6

VALIDATION OF EXISTING MODEL CANTY

6.1 INTRODUCTION

Model validation is a procedure that determines how closely a model's behaviour fits the real world and to what extent logically and biologically the model agrees with actual forest growth (Goulding, 1979; Vanclay, 1994; Soares *et al.*, 1995). Model CANTY was built by the New Zealand Forest Research Institute (NZ FRI) and it is for radiata pine growing in Canterbury (Goulding, 1995). The objective of the validation of model CANTY in this study was to determine how closely the model predicted stand growth in SPBL's estate. The source and magnitude of discrepancies between predictions and observations needed to be identified.

The model can be examined quantitatively and qualitatively. Quantitative validation aims to assess model behaviour with a set of data, and this procedure involves graphical displays and statistical tests. Inspection of graphical plots of residuals versus predictions or potential explanatory variables is an efficient way to detect potential correlation. Residual errors may display certain trends along with initial conditions, projection length, or predicted values when the model is biased. Linear regression of residuals versus explanatory variables is a statistically efficient way to detect the magnitude of linear correlation (Soares *et al.*, 1995).

Among several measures describing model performance, average model bias (AMB) and model efficiency (EF) (Loague and Green, 1991) are widely used. Average model bias (Equation 6.1) is an average of errors (mean residuals) for overall predictions. A value of zero of AMB is expected for a good model. EF (Equation 6.2) is also a measure for overall performance and it was recommended by Loague and Green (1991) and Mayer (1993). A high value of EF up to one is a perfect fit, zero means a prediction equivalent to mean value, and a negative value indicates a worse prediction.

a) average model bias $$AMB = \tfrac{1}{n}\Sigma(Y_i - \hat{Y}_i) \tag{6.1}$$

b) model efficiency or efficiency factor $$EF = 1 - \frac{\Sigma(Y_i - \hat{Y}_i)^2}{\Sigma(Y_i - \bar{\bar{Y}}_i)^2} \tag{6.2}$$

where Y_i = observed value, $\hat{Y}_i$ = predicted value, n = sample size

Data used for this validation were all plot measurements from SPBL's estate that represented its radiata pine stands. They covered a wide range of stand conditions described in Chapter 3. The data were independent of the data used in building CANTY. The model could fit very well to plot measurements with which it was built but it might or might not work completely with this independent data set because of the distinctive silvicultural regime that SPBL practises and the different site conditions in SPBL's estate.

Qualitative validation examines the biological aspects of each single module and the logical structure of the model as a whole. The model CANTY is a "state space" system and the theoretical aspects of the whole structure have been well documented by Garcia (1984, 1988, 1994). Only a simple description of the system follows.

6.2 The Existing Model in Canterbury

The existing model CANTY is a state-space model. In a state-space system, a state is defined by several state vectors. A state of a stand in NZ FRI models is usually expressed by mean top height, basal area, stems per hectare, and sometimes crown closure. Future states can be predicted by future management options and current states that are historical summaries of the past. The behaviour of the system is described by a transition function and an output function. The output function for volume is predicted with state variables. The transition function should possess the properties of consistency for zero elapsed time, the equity of final projection from different steps, and the invariability of two states if inputs are the same for a given period and a given initial state. Integration of differential equations or differentiation of difference equations can satisfy the above properties automatically (Garcia, 1984; Garcia, 1988; Garcia, 1994).

Some theoretical aspects of the model formulation are:

Transition function: $\mathbf{X(t) = F[X(t_0),\ \ U,\ \ T - T_0]}$ or $\mathbf{dX/dt = f(X)}$

Output Function: $\mathbf{V(t) = G[X(t)]}$, or $\mathbf{dV/dt = g(X)}$

where $\mathbf{X = (H, G, N)}$, H = mean top height, G = basal area, N = stocking

V = Volume, U = input (management options)

T_0 = starting time of a period, T = the end of a period of time

When a multivariate generalisation of the Bertlanffy-Richards model is adopted, the new state vectors become:

$\mathbf{Y = (y_1, y_2, y_3) = (H^{c_{11}}, H^{c_{21}} G^{c_{22}} N^{c_{23}}, H^{c_{31}} G^{c_{32}} N^{c_{33}})}$ or $\mathbf{Y = X^c}$.

The linear differential equation is:

$\mathbf{dY/dt = AY + b = AX^c + b}$ or

$\mathbf{dY/dt = A(X^c - a)}$ when $\mathbf{a = -A^{-1}\,b}$

where A and C is 3×3 matrix, a and b is 1×3 matrix, T = $b_h t$ is scaled time, b_h is the coefficient in site index equation.

A global difference equation $\mathbf{X(t_2) = \{a + P^{-1} e^{\Lambda b_h (t_2 - t_1)} P\ [X^c(t_1) - a]\}^{(1/c)}}$ is obtained from integration of the above differential equation. P and Λ are such that Λ is diagonal and $\mathbf{A = P^{-1}\,\Lambda\,P}$.

A maximum likelihood estimator was used to estimate all equations simultaneously and the overall errors were minimised in the system. A maximum likelihood estimator might be theoretically superior to least-squares estimation methods, but the differences in values of estimated parameters were not always of practical importance as shown in an early study by Sullivan and Clutter (1972). The simultaneous estimation of all model components minimises overall errors but restricts the choice of functional form for individual state variables.

The existing model CANTY is a stand level model including the components of mean top height, basal area/ha, stems/ha and volume/ha. Juvenile growth of basal area was projected with given stocking and mean top height (typically from 4 m to 8 m). No height and stocking equations were available for juvenile growth. All equations were created with pooled data from foothill and plains plantations.

6.3 METHODS

Model examination procedures in this study employed statistical regression of residuals against predictions and against explanatory variables. Graphical plotting of residuals against predictions and against explanatory variables was included. Such measures of model performance as AMB (average model bias) and EF (efficiency factor) were calculated with a full data set.

To test the significance of a number of variables with a regression procedure, a data set comprising 191 autocorrelation-free plot measurements was used. Hypothesis tests based on autocorrelated data are invalid (Neter and Wasserman, 1974; West *et al.*, 1984; West, 1995). Each observation was a random pair of the repeated plot measurements in each stand on the plains and a random pair in each plot on the hills as all plot altitudes were the same in a stand on the plains.

The tested variables (the main components of the model) were mean top height (MTH), basal area/ha (G), stems/ha (N) and volume/ha (V). The explanatory variables were age, time increment (projection length), geographic type, altitude, and GF rating. Geographic type was a binary variable (0, 1) representing plains and hills, respectively.

6.4 RESULTS

Predictions were calculated with model CANTY and residuals were derived from the actual minus the predicted values. Using an autocorrelation-free data set, regression results for residuals of mean top height, basal area/ha, stems/ha, and volume/ha are shown in Tables 6.1 to 6.4. In each table, regression results for both residuals versus predictions and residuals versus explanatory variables are listed.

Graphical plots are shown in Figures 6.1 to 6.4, using the full data set. In each figure, the four jointed plots were residuals versus predicted value, age, altitude and time increment (projection length), respectively.

Average model bias and efficiency factor were calculated for each model component and are listed in Table 6.5.

On the whole the model CANTY was biased for predicting production of stands in SPBL. The specifics for each model component were summarised as follows.

(1) The regression results of residuals of MTH against prediction showed that both intercept and slope were significantly different from zero (Table 6.1). The results of regression of residual versus explanatory variables indicated in the same table that residuals showed significant trends with projection length (time increment) and altitude. The model overestimated MTH with a negative average model bias (AMB) of -1.059 m for all plots (Table 6.5). Bias was more serious at lower altitudes and with longer projection lengths, and up to 6.0 m overestimates were displayed. The regression numerical results were in accordance with the residual pattern shown in Figure 6.1. The overestimation for MTH in the database indicated that height growth rate in SPBL's estate was lower than that of PSP measurements

with which CANTY was built. Site index was not included in regression analyses for all dependent variables, as it must be biased when the MTH projection model is biased.

Table 6.1: Results of regression analysis of residuals of MTH

Dependent Variable	Independent Variable	DF	Parameter Estimate	Std. Error	T	Prob > |T|
Residual of MTH	Intercept	1	1.0244	0.3708	2.763	0.01
(R^2= 0.1345)	Prediction	1	-0.1075	0.0198	-5.419	0.0001
	Intercept	1	-0.4148	0.4310	-0.962	0.3371
	Age	1	0.0249	0.0214	1.161	0.2473
Residual of MTH	Time Increment	1	-0.2919	0.0287	-10.164	0.0001
(R^2= 0.4535)	Geography type	1	-0.1617	0.3430	-0.472	0.6378
	GF rating	1	-0.0412	0.0259	-1.591	0.1133
	Altitude	1	0.0042	0.0011	3.996	0.0001

(2) The regression results of residual of basal area/ha (G) against prediction showed that both intercept and slope were not significantly different from zero (Table 6.2). There was, therefore, no significant bias for basal area as a whole. The results of regression of residual versus explanatory variables indicated that residuals showed significant trends with altitude, a clear trend which can be seen with Figure 6.2. The trend existed within both plains and foothills. It could not be described by a simple indication of geographical type which was not a significant variable when altitude was included in the regression model. The model overestimated basal area seriously for low altitudes on the plains and underestimated more seriously for higher altitudes on the hills. A negative -1.561 of AMB was obtained for plains and a positive 6.276 of AMB was shown for the hills (seen in Table 6.5). EF reduced to 0.75 for the hills. A possible hypothesis worth examining was that the model could be improved by including altitude in the projection model of basal area and mean top height.

Table 6.2: Results of regression analysis of residuals of G

Dependent Variable	Independent Variable	DF	Parameter Estimate	Std. Error	T	Prob > |T|
Residual of G	Intercept	1	-0.8018	0.9647	-0.831	0.407
(R^2= 0.0057)	Prediction	1	0.0253	0.0243	1.04	0.2999
	Intercept	1	-4.5462	1.8403	-2.47	0.0144
	Age	1	-0.0545	0.0915	-0.595	0.5523
Residual of G	Time Increment	1	0.0282	0.1226	0.23	0.8186
(R^2= 0.3587)	Geography type	1	-2.8859	1.4646	-1.97	0.0503
	GF rating	1	-0.0178	0.1105	-0.161	0.8722
	Altitude	1	0.0303	0.0045	6.694	0.0001

(3) Table 6.3 shows that both intercept and slope were not significantly different from zero in the regression of residual of stocking against prediction. The results of regression of residual versus explanatory variables indicated that residuals showed marginal significance with initial age and GF rating although the trends of residuals with initial age cannot be seen visually on Figure 6.3. The fitting difference between plains and hills cannot be seen visually either. The least bias of the model CANTY was for projections of stocking/ha.

(4) The projection performance for volume was similar to that for basal area. The regression results of residual of volume against prediction showed that both intercept and slope were not significantly different from zero (Table 6.4). The results of regression of residual versus explanatory variables indicated significant trends with altitude and time increment. A clear trend of altitude can be seen with Figure 6.4. The model produced more serious bias with overestimates for lower altitudes on the plains and with underestimates for higher altitudes on the hills. The underestimates for hills were not negligible; the magnitude can be seen clearly with Figure 6.4 and was confirmed by the second lowest EF value 0.816 in Table 6.5. A difference existed not only between geographical type (plains and foothills) but also within one type with altitude.

Table 6.3: Results of regression analysis of residuals of N

Dependent Variable	Independent Variable	DF	Parameter Estimate	Std. Error	T	Prob > |T|
Residual of N	Intercept	1	-6.8264	9.0519	-0.754	0.4517
(R^2= 0.0902)	Prediction	1	0.0019	0.0108	0.177	0.86
	Intercept	1	-47.2613	20.5522	-2.3	0.0226
	Age	1	2.3306	1.0221	2.28	0.0237
Residual of N	Time Increment	1	2.1662	1.3694	1.582	0.1154
(R^2= 0.0936)	Geography type	1	-23.3489	16.3557	-1.428	0.1551
	GF rating	1	2.6245	1.2340	2.127	0.0348
	Altitude	1	-0.0181	0.0505	-0.358	0.7209

Table 6.4: Results of regression analysis of residuals of V

Dependent Variable	Independent Variable	DF	Parameter Estimate	Std. Error	T	Prob > |T|
Residual of V	Intercept	1	1.7009	6.3904	0.266	0.7904
(R^2=0.0108)	Prediction	1	-0.0323	0.0225	-1.433	0.1534
	Intercept	1	-14.9080	15.7886	-0.944	0.3463
	Age	1	-0.9441	0.7852	-1.202	0.2307
Residual of V	Time Increment	1	-3.5653	1.0520	-3.389	0.0009
(R^2=0.3449)	Geography type	1	-20.3909	12.5648	-1.623	0.1063
	GF rating	1	-1.1287	0.9480	-1.191	0.2353
	Altitude	1	0.2400	0.0388	6.189	0.0001

Table 6.5: The results of model performance measures

Model	Overall		Plains		Foothills	
component	AMB	EF	AMB	EF	AMB	EF
MTH	-1.059	0.825	-1.112	0.810	-0.279	0.889
G	-1.068	0.925	-1.561	0.935	6.276	0.751
N	-2.170	0.968	-1.299	0.962	-15.12	0.979
V	-14.68	0.915	-18.29	0.913	38.98	0.816

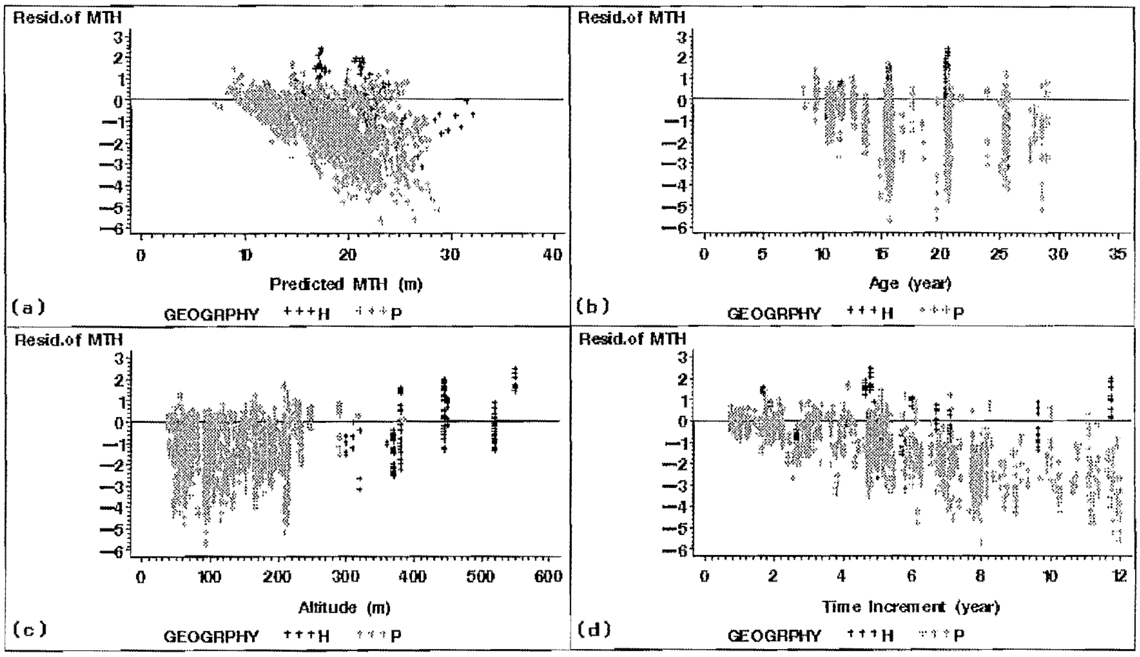


Figure 6.1: The residual pattern for fit of MTH

(a) Residual vs. predicted value; (b) Residual vs. initial age; (c) Residual vs. altitude; (d) Residual vs. time increment (projection length)

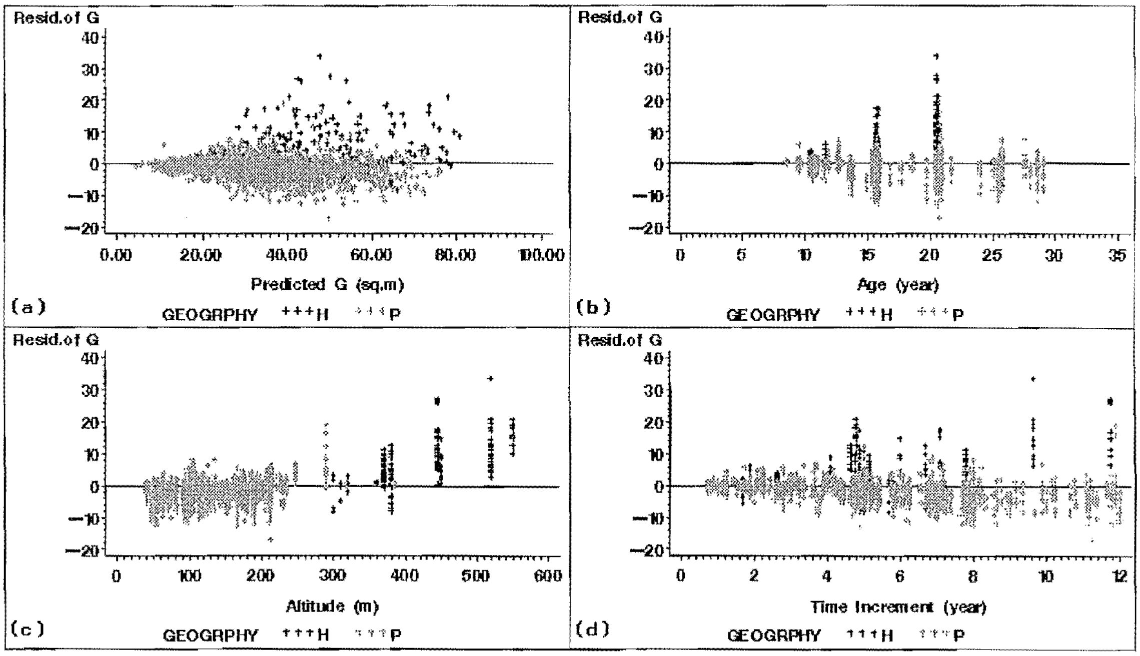


Figure 6.2: The residual pattern for fit of basal area per hectare

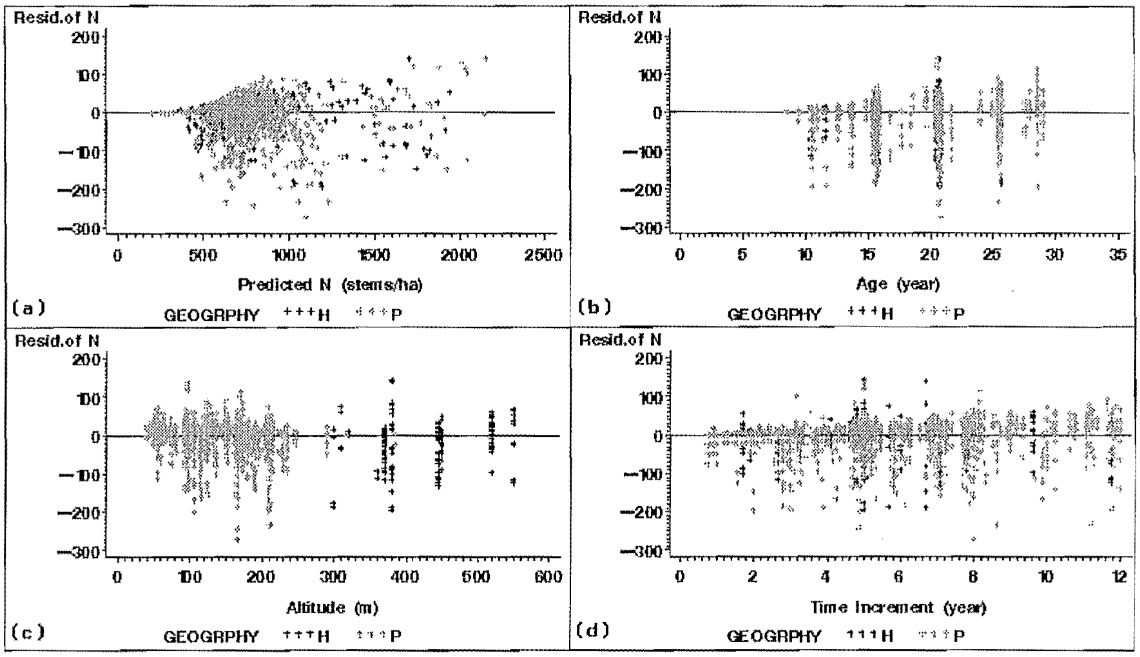


Figure 6.3: The residual pattern for fit of stems per hectare

(a) Residual vs. predicted value; (b) Residual vs. initial age; (c) Residual vs. altitude; (d) Residual vs. time increment (projection length)

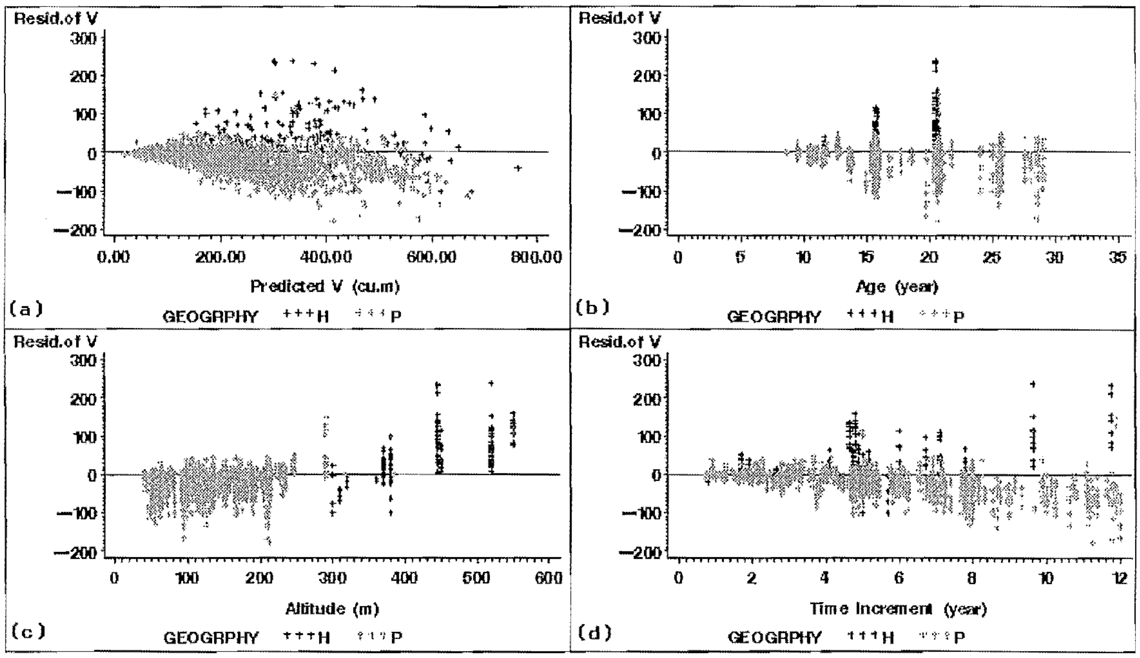


Figure 6.4: The residual pattern for fit of volume per hectare

6.5 CONCLUSIONS

Regression analyses and graphical procedures revealed that the model CANTY overestimated MTH by 1 m on average over the whole SPBL's estate. The residuals of MTH, basal area/ha and volume/ha were significantly related to altitude. The model produced more serious overestimates for lower altitudes on the Canterbury plains and underestimates for higher altitudes on hill sites. No significant bias was detected for stocking projections.

CHAPTER 7

DEVELOPMENT OF MODEL CanSPBL (1): STAND MODEL

7.1 INTRODUCTION

No single stand model could be expected to provide enough information for all levels of decision-making (Burkhart, 1977; Leary, 1979; Daniels and Burkhart, 1988). Stand models (named whole stand models by Munro, 1973) can be used for general and long-term planning while individual tree models can be used for detailed planning.

A new model, CanSPBL, was developed to represent both stands and individual trees in SPBL's estate. Chapter 7 describes the establishment of the model at a stand level and the following chapter describes how the model was extended to allow projections at tree level. The components of model CanSPBL at the stand level comprised mean top height (MTH), basal area per hectare, stems per hectare, volume per hectare and diameter distribution.

Improving model projection accuracy is a constant subject for research. Modellers have proposed and tested various approaches. To increase reliability of model CanSPBL, the following aspects were given particular attention:

(1) A good sample data set covering a target population spatially and temporally (Vanclay 1994) is necessary for constructing a model. Serious bias may result from a database with inappropriate coverage. CanSPBL is a model for SPBL's estate and for whole rotations. Samples have been obtained covering a wide range of conditions as described in Chapters 3 and 5.

(2) A method of preparing data structure to deal particularly with the time increments (or growth interval), should be considered in order to improve long-term projection ability. Managers often need longer-term predictions of yield than the originally scheduled time intervals of re-measurement, but it is often unknown whether or not the projection ability is robust enough for a ten year-long projection from a model produced with measurements of two-year increments. An effective way to ensure longer-term prediction ability is to include longer-term measurements when creating a model (Lee, 1998). Borders *et al.* (1987) investigated the model performance of three methods of data preparation --- non-overlapping growth intervals (NI), all possible growth intervals (AI), and longest intervals (LI). It was concluded that projection ability was model form dependent. Further studies are needed.

(3) Sigmoid difference equations are commonly employed for yield projections. Different equation forms behave differently (Woollons and Wood, 1992), so appropriate equation forms was selected after testing a large variety of equations.

(4) The incorporation of significant environmental variables can improve model projection ability. A difference equation is an efficient projection form because the initial values of dependent variables represent past history of growth and account for great proportions of growth differences in a period. Some environmental variables, however, might still improve projection performance (Woollons *et al.*, 1997, for example). Explanatory variables that are easily available to managers should be tested and incorporated in models if they make projections significantly more precise.

(5) Appropriate use of a regression estimator can create a statistically stable model. West (1995) proposed fitting a model with a full data set and testing it with an autocorrelation-free data set, when least-squares methods are applied to repeated measurements of permanent sample plots.

A stand model can be built with statistics at a plot or stand level. Plot statistics, as described in Chapter 3, were used in this study to develop the model components of MTH, basal area, stems per hectare and volume per hectare while stand level statistics of diameter deviation and maximum diameter were used to fit reverse Weibull distributions.

For describing size-class distribution of even-aged plantations in New Zealand, the Weibull distribution function has commonly been used because of its flexibility, simplicity, and reliability (Whyte and Woollons, 1992). Maximum diameter employed in a reverse Weibull distribution has resulted in a better fit because maximum diameter could be projected more reliably than minimum diameter (Kuru *et al.*, 1992; Xu *et al.*, 1992).

A problem of distribution models is the spatial correlation of tree diameters when predictions for stands are considered. Such stand measures from a random plot as mean height, basal area/ha, stems/ha, and volume/ha are not biased but variance, percentiles and extreme values may vary with plot size (Garcia, 1991). Extreme value percentiles have been used to estimate location parameters in reverse Weibull distributions (Xu *et al.* 1992) on a basis of plot rather than stand level. This study used different approaches to estimate standard deviation and maximum diameter at a stand level. Standard deviation of a stand was obtained from the cluster sampling formulation (Garcia, 1991). The maximum diameter was estimated from all plots in a stand.

The newly established model was validated. Two sources of data at a stand level and a plot level were used to test model projections of MTH, basal area per hectare, stems per hectare, and estimations of volume per hectare.

7.2 METHODS AND PROCEDURES

To fit equations, non-linear least-square procedures with SAS software (SAS Institute Inc., 1992) were employed and the Mean Square Error (MSE) and graphical residual patterns were used as the selection criteria to judge model performance. Skewness and kurtosis were often checked for final models to determine the magnitude of residual distributions from normality. Plots of residuals versus predictions and all possible explanatory variables were inspected to check for trends but only the four most important graphs are displayed here for each selected equation. The four graphs are residuals versus prediction, age, altitude and time increment. Age was the main independent variable in equations and projection performance should be consistently balanced with time. The effects of altitude had been revealed during tests of model CANTY. Plots of residuals versus projection length (time increment) were the final graphs to show the lack of bias over long projections.

To test the significance of explanatory variables for a model, ordinary regression analyses of residuals were conducted, always with a set of autocorrelation-free data. To test altitude efficiently, an independent data set was prepared in such a way that only one pair of the repeated plot measurements was chosen at random from each stand on the plains and a pair from each plot on the hills. Altitude was almost the same throughout each individual stand on the plains. Data with autocorrelation due to repeated measurements create unbiased coefficients with least-squares regression, but underestimate both the variances of error terms and the variances of coefficients. In other words, hypothesis tests based on autocorrelated data are invalid (Neter and Wasserman, 1974; West *et al.*, 1984; West, 1995) but there are no problems with parameter estimates obtained by fitting models to such data.

To describe diameter-class distributions at a stand level, a reverse Weibull function was employed. To produce stand tables, the recovery method was used to project future from the current stand statistics and the method of moments was used to convert the stand statistics of standard deviation, maximum value, and arithmetic mean diameter to Weibull distribution parameters. Standard deviation of a stand was obtained from the cluster sampling formulation used by Garcia (1991). Maximum diameter was estimated from all plots in a stand. Projection equations for standard deviation and maximum diameter were created. Arithmetic mean diameter was derived from the outputs of projection equations of basal area and stocking.

To validate the newly established model and test its performance for use at both a stand and a plot level, the same methods were used as for validation of model CANTY. The main model components of MTH, basal area per hectare, stems per hectare, and volume per hectare were examined using two sources of data at a stand and a plot level. One source of data was a set of independent plot measurements chosen randomly before building the model. The other was the same as data used in the model establishment but reorganised for stand level analysis. The average model bias (AMB), efficiency factor (EF), and skewness were calculated for each model component. Graphs of residual patterns were examined to detect bias. Regressions of residuals versus both explanatory variables and predictions were separately conducted with autocorrelation-free data.

7.3 Analyses and Results

7.3.1 Modelling of mean top height and basal area per hectare

To model the growth of MTH and basal area/ha, various equations were tried, ways of preparing data to increase projection ability for long projection were examined, and altitude effects were formulated into the selected model. Residual regression analyses were conducted, revealing that no further explanatory variables could be found to include in the model. These steps are explained in detail in this section.

7.3.1.1 Model fitting and selection from difference equations

Stand basal area and mean top height usually display a sigmoid shape over a whole rotation. Sigmoid equations proposed for modelling growth in forestry include the log-reciprocal Schumacher model (Schumacher, 1939; Clutter, 1963), Chapman-Richards model (Von Bertalanffy, 1949; Richards, 1959; Pienaar and Turnbull, 1973), Weibull model (Yang *et al.*, 1978), Gompertz (Nokoe, 1978), Hossfeld (Woollons *et al.*, 1990), and others (Woollons and Wood, 1992) such as logistic and mono-molecular. Anamorphic and polymorphic forms were identified according to whether or not a proportionality relationship holds within a curve family.

These commonly used difference equations were fitted during this study for MTH and basal area per hectare. Data described in Chapter 3 were reorganised for fitting projection equations. A total of 2692 observations were included for fitting MTH, and 2593 observations were used for fitting basal area/ha. Table 7.1 lists the Mean Square of Error (MSE) for both MTH and basal area for the different equation forms. For MTH, a three-parameter polymorphic Weibull equation was almost not different in MSE from a two-parameter polymorphic Schumacher equation. The two-parameter polymorphic Schumacher equation also led to the smallest MSE for basal area. The polymorphic Schumacher form with two parameters was therefore chosen as the basic form for both MTH and basal area. Graphical patterns of residuals from this equation are shown in the later sections.

7.3.1.2 The determination of time increment (measurement interval or projection length)

An effective way to improve prediction ability, longer-term projections in particular, is to include longer-term measurements when creating a model (Lee, 1998). If there are three measurements in a permanent sample plot at T1, T2, and T3, the traditional way of arranging records is to link two adjacent measurements (non-overlapping short intervals), T1 to T2, T2 to T3. A data structures of all possible intervals (AI) created by joining any two measurements will result in T1 to T2, T2 to T3, and T1 to T3. Quite a number of longer time increments can be created using this latter approach, although the number can theoretically be smaller than the number of short intervals.

Table 7.1: Difference equation forms and fitting results for MTH and basal area

Difference equation Forms	MSE	
	MTH	G
Polymorphic Schumacher I: $Y_2 = \exp[\ln(Y_1)(T_1/T_2) + \alpha(1-(T_1/T_2))]$	0.929	10.399
Polymorphic Schumacher II: $Y_2 = \exp[\ln(Y_1)(T_1/T_2)^\beta + \alpha(1-(T_1/T_2)^\beta)]$	0.697	10.375
Anamorphic Schumacher: $Y_2 = Y_1 \exp[-\beta(\frac{1}{T_2^\gamma} - \frac{1}{T_1^\gamma})]$	0.861	16.085
Polymorphic monomolecular: $Y_2 = Y_1 \exp(-\beta(T_2 - T_1)) + \alpha[1 - \exp(-\beta(T_2 - T_1))]$	0.846	14.610
Anamorphic monomolecular: $Y_2 = Y_1[1-\beta\exp(-\gamma T_2)]/[1-\beta\exp(-\gamma T_1)]$	0.860	17.056
Polymorphic Levakovic: $Y_2 = [Y_1^\gamma (T_1/T_2)^2 + \alpha(1-(T_1/T_2)^2)]^{1/\gamma}$	0.728	11.637
Anamorphic Levakovic: $Y_2 = Y_1[(T_2/T_1)^2(\beta + T_1^2)/(\beta + T_2^2)]^\gamma$	0.859	15.753
Polymorphic Hossfeld: $Y_2 = [(1/Y_1)(T_1/T_2)^\beta + \alpha(1-(T_1/T_2)^\beta)]^{-1}$	0.715	12.362
Anamorphic Weibull: $Y_2 = Y_1[1-\exp(-\beta T_2^\gamma)]/[1-\exp(-\beta T_1^\gamma)]$	0.858	16.218
Polymorphic Weibull: $Y_2 = \alpha - \beta[(\alpha - Y_1)/\beta]^{(T_2/T_1)^\gamma}$	0.696	10.655
Anamorphic Gompertz: $Y_2 = Y_1 \exp[-\beta\exp(-\gamma T_2)]/\exp[-\beta\exp(-\gamma T_1)]$	0.889	15.926
Polymorphic Gompertz I: $Y_2 = \exp[\ln(Y_1)\exp(-\beta(T_2 - T_1))]\exp[\alpha(1-\exp(-\beta(T_2 - T_1)))]$	0.900	16.059
Polymorphic Gompertz II: $Y_2 = \exp[\ln(Y_1)\exp(-\beta(T_2 - T_1) + \gamma(T_2^2 - T_1^2))]\exp[\alpha(1-\exp(-\beta(T_2 - T_1) + \gamma(T_2^2 - T_1^2)))]$	0.745	10.735
Anamorphic Chapman-Richards: $Y_2 = Y_1[(1-\exp(-\beta T_2))/(1-\exp(-\beta T_1))]^\gamma$	0.857	15.766
Polymorphic Chapman-Richards: $Y_2 = [\alpha - (\alpha - Y_1^{1/\gamma})\exp(-\beta(T_2 - T_1))]^\gamma$	0.844	14.564

Note: Y = MTH or G, T = age, α, β and γ = parameters

There are advantages and disadvantages with the inclusion of long-term intervals. When measurement errors and annual weather fluctuations occur, yield can be estimated more precisely with longer time increments. Measurement errors in relative terms tend to increase with short time increments (Vanclay, 1994). Figure 7.1 shows that the variation of estimated mean annual increment (MAI, real mean annual increment confounded possibly with measurement error) of mean top height became smaller when projection length increased. The largest variations occurred when time increment was less than three years. Autocorrelation is more serious with overlapped, all possible intervals, however. Tests with least-squares estimation can be validly done only by using a set of extracted data comprising randomly selected single observations for each plot.

In SPBL, the repeated inventory was scheduled every 3 to 5 years but quite a number of plots were measured annually.

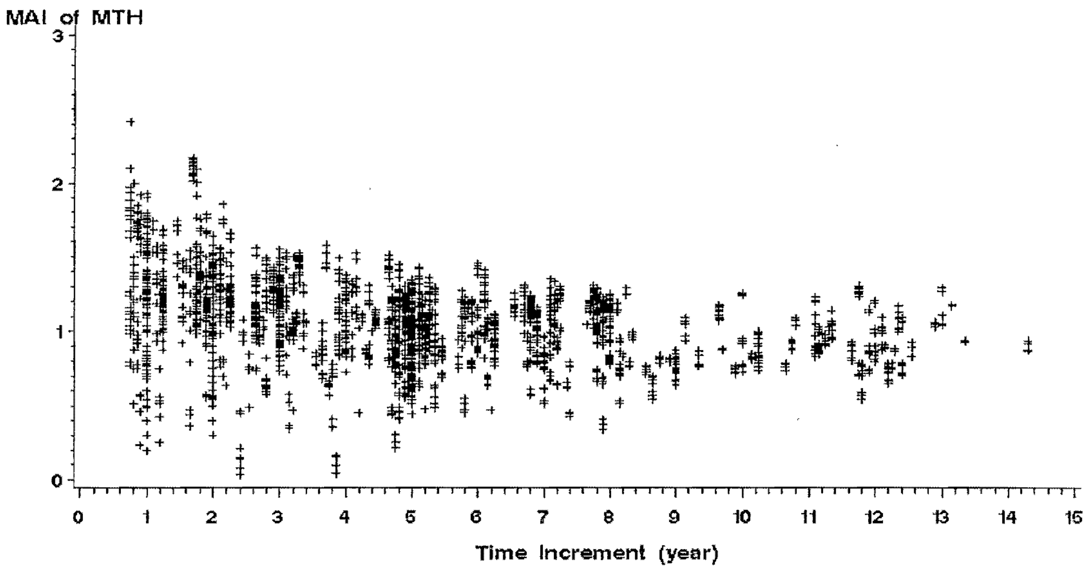


Figure 7.1: Variation of mean annual increment of height decreased with time increment

In this study, a null hypothesis was that the long-period prediction capability of a model created with the inclusion of long-period measurements was the same as that of a model created without this inclusion. Testing the hypothesis was demonstrated in terms of MTH.

Five sets of data were prepared for the tests. The time increment frequency in each data set is listed in Table 7.2. The first set comprised data with short time increments from

0.7 to 4 years, the second from 0.7 to 5 years, the third from 0.7 to 6 years, the fourth from 0.7 to 12 years. The fifth set comprised only long time increments from 6 to 12 years. The first up to fourth sets were used for the establishment of each model, and the last set was used to test and compare the four models. A few extraordinary observations with time increments longer than 12 years or shorter than 0.7 years were not included due to their small frequency and large variation.

Table 7.2: Time increment frequency in each data set

Time Increment	1	2	3	4	5	6	7	8	9-12	Sum
Data set 1 Freq.	77	370	335	251	0	0	0	0	0	1033
Data set 2 Freq.	77	370	335	251	373	0	0	0	0	1406
Data set 3 Freq.	77	370	335	251	373	448	0	0	0	1852
Data set 4 Freq.	77	370	335	251	373	448	242	273	323	2692
Data set 5 Freq.	0	0	0	0	0	448	242	273	323	1286

The same form of a 2-parameter polymorphic Schumacher equation was fitted separately for the above four data sets and each of the fitted four models was applied to the test data set. Table 7.3 lists the model residual statistics for each model. Figure 7.2 shows the residual pattern versus prediction for the four models. Graph (a) in Figure 7.2 shows an overall overestimation but others show little. Mean residuals from the first to the fourth data sets were -0.5499, -0.2711, -0.2280 and -0.0503, respectively.

Table 7.3: Residual statistics for models built with different data structure

Model	N	Mean	Std dev	Minimum	Maximum
Model 1 (with data set 1)	1286	-0.5499	1.1744	-7.5872	2.3943
Model 2 (with data set 2)	1286	-0.2711	0.8852	-3.3332	2.6011
Model 3 (with data set 3)	1286	-0.2280	0.8863	-3.2290	2.5221
Model 4 (with data set 4)	1286	-0.0503	0.8666	-3.0959	2.7484

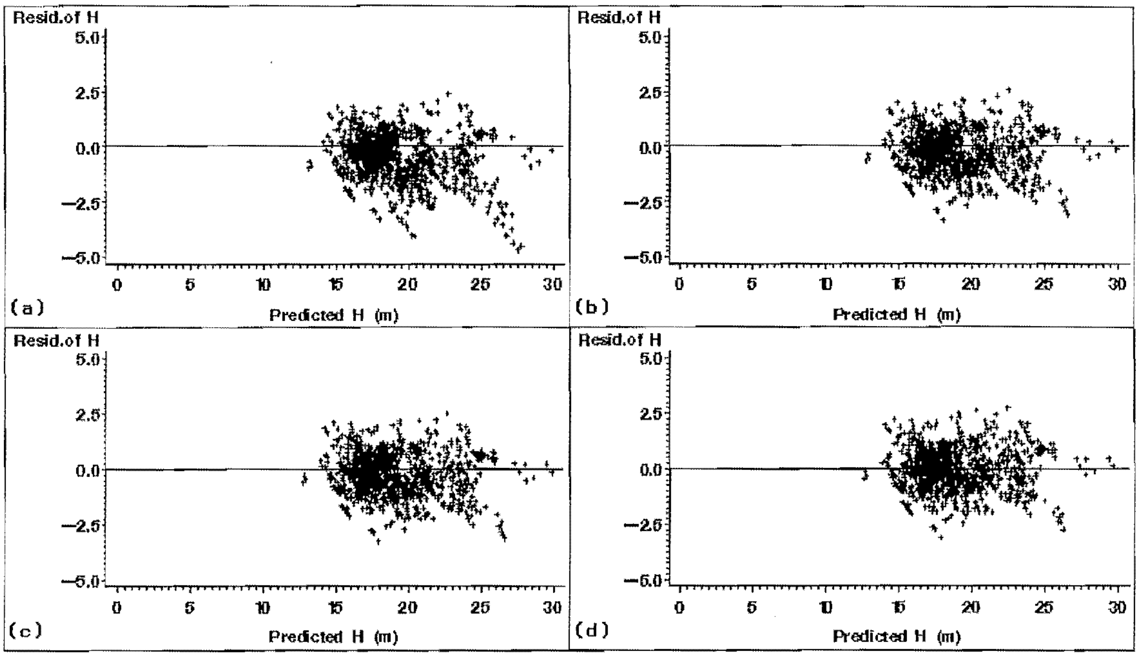


Figure 7.2: Residual pattern for longer-term projections with models built of different data structures

(a) The residual pattern of model 1 created with data set 1

(b) The residual pattern of model 2 created with data set 2

(c) The residual pattern of model 3 created with data set 3

(d) The residual pattern of model 4 created with data set 4 (interpolation)

Table 7.4 lists the results of regression tests of residuals versus prediction with a data set in which one observation was chosen randomly from each plot in data set 5. It is shown that the intercept and slope were significantly different from zero with model 1 (R^2=0.33), while they were not significantly different from zero with other models. Negative correlation (negative slope) was displayed between the residuals and predictions with model 1.

It could be concluded that among four models, model 1 was significantly biased but others were not. It can also be seen that model 4 is not proved statistically different from model 2 and model 3 but it tended to lead to a better fit for long projections in terms of the value of mean residual and less bias shown in the graphical residual pattern. A data structure equivalent to that of data set 4 (with the all-possible intervals) was selected in this study for fitting projection models because it seemed that although it could not be proven statistically to be better than all others, neither would it provide a poorer fit.

A few plots having time increments longer than four years were dropped from data set 1, because the actual time intervals of re-measurements were scheduled at 3 to 5 years but data set 1 included records with maximum time increments of 4 years only. The partial loss of information caused the model built with data set 1 to be significantly biased for long projections. Projections were less biased; and performances were similar using models created with data sets 2, 3 and 4 in which the maximum time increments were larger than or equivalent to the maximum of actual re-measurement intervals. It was probably the representative ability of the data sets that made the projections different or similar. The equations used have the desirable property of path-invariance.

Table 7.4: Regression test results for residuals from different models

Model	Variable	DF	Parameter Estimate	Std Error	T	Prob >|T|
Model 1	Intercept	1	4.540723	0.6711	6.766	0.0001
	Prediction	1	-0.28586	0.032011	-8.93	0.0001
Model 2	Intercept	1	-0.1680	0.4658	-0.361	0.7189
	Prediction	1	-0.0145	0.0234	-0.619	0.5369
Model 3	Intercept	1	0.0114	0.4572	0.025	0.9801
	Prediction	1	-0.0200	0.0230	-0.871	0.3852
Model 4	Intercept	1	-0.3713	0.4669	-0.795	0.4276
	Prediction	1	0.0097	0.0237	0.410	0.6825
	Error	158				

Borders *et al.* (1987) found that a data structure of non-overlapping short intervals was best for one equation form and a data structure of long intervals was best for the other.

It was noted that projection ability with different data structures depended on the forms employed. This study suggests that projection ability might be data source dependent. Lee (1998) demonstrated all possible mixed intervals (AI) was the best data structure in his study of growth and yield modelling for Douglas fir plantations in the South Island of New Zealand. A theoretical study may be needed to clarify the issue. This study was not specifically focused on finding a universal solution for a theoretical problem, but rather adopting a best data structure for model development.

7.3.1.3 Calibration of the effect of altitude

As Figure 7.3 and Figure 7.4 show clearly, increasing trends of residuals of both MTH and basal area/ha were found with altitude. A straight line could not describe the effect well, however. The slope of the pattern for the hills was sharper than that for the plains. The magnitude of the pattern for basal area was more obvious than for MTH. To calibrate the effect of altitude on the growth of MTH and basal area, the model had to account for growth differences within and between plains and hills.

It seemed that residual pattern changed direction at altitudes around 250 m, at which the geographical difference of plains and hills is mostly identified. Based on that assumption, a regression test model for the residuals was formulated (Equation 7.1).

$$Residual = \alpha_0 + \alpha_1\, Altitude + \alpha_2\, (Altitude - 250)\, X \qquad (7.1)$$

$$\begin{cases} Residual = \alpha_0 + \alpha_1\, Altitude & \text{when X=0 (altitude<250)} \\ Residual = (\alpha_0 - 250\, \alpha_2) + (\alpha_1 + \alpha_2)\, Altitude & \text{when X=1 (altitude}\geq\text{250)} \end{cases}$$

where Residual = residual in the model of MTH or basal area

α_0, α_1, α_2 = coefficients

X = binary indicator variable, X=0 if altitude<250 and X=1 if altitude≥250.

A regression analysis of residuals against altitude was done to test the altitude effect with a set of independent, autocorrelation-free data. It was shown (Table 7.5) that altitude was significant for both MTH and basal area/ha on plains and foothills.

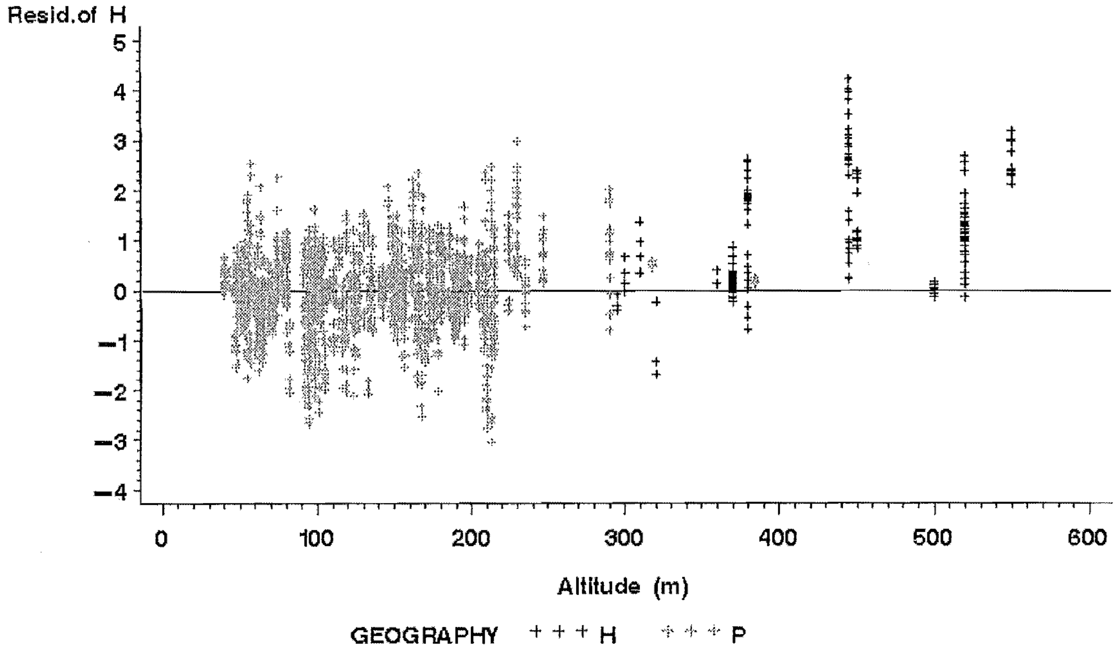


Figure 7.3: The trends of residual of MTH versus altitude

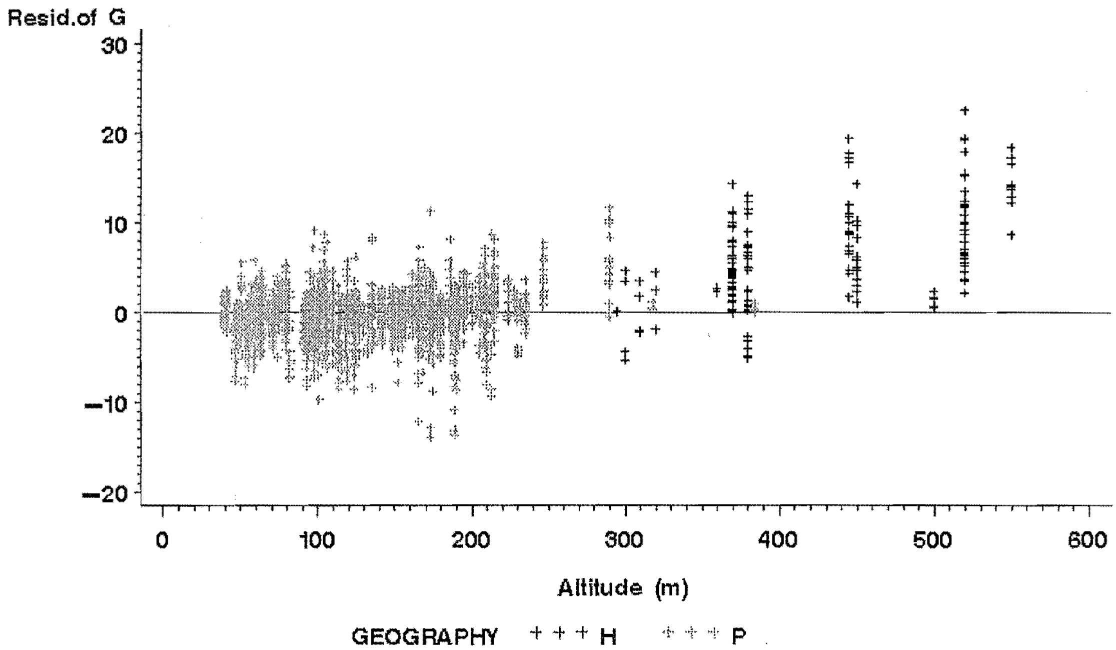


Figure 7.4: The trends of residual of basal area versus altitude

Table 7.5: Regression tests of residuals of MTH and basal area versus altitude

Variable	Residuals for MTH				Residuals for G			
	Alt<250 Total df=142		Alt≥250 Total df=60		Alt<250 Total df=142		Alt≥250 Total df=60	
	Parameter Estimate	Prob>|T|	Parameter Estimate	Prob>|T|	Parameter Estimate	Prob>|T|	Parameter Estimate	Prob>|T|
Intercept	-0.5348	0.0021	-1.2030	0.0881	-2.0550	0.0006	-9.9029	0.0075
Altitude	0.0032	0.0095	0.0054	0.002	0.0140	0.0012	0.0397	0.0001

To transfer the effect of altitude into the original Schumacher equation, the asymptotic parameter in the Schumacher equation was replaced by the same form as Equation 7.1, and then the full equation was transformed into Equation 7.2.

$$Y_2 = \exp[\ln(Y_1)(T_1/T_2)^{\beta} + (\alpha_0 + \alpha_1 Alt + \alpha_2 (Alt-250)X)(1-(T_1/T_2)^{\beta})] \quad (7.2)$$

where Y_2 is future mean top height or basal area/ha, Y_1 is the initial mean top height or basal area, T_1 is initial age and T_2 is the future age, Alt is altitude, β is parameter, α_0, α_1, α_2 and X are the same as the above.

Non-linear procedures were run again after the inclusion of altitude. Table 7.6 lists the results of MSE, standard deviation, skewness and kurtosis before and after the inclusion of altitude in the model of MTH and basal area. Altitude improved the model, with reductions of MSE of 17% for MTH and 41% for basal area. There were also large reductions in standard deviation, skewness and kurtosis. The parameter estimates of the final models are listed in Table 7.7. The model showed no signs of bias for both MTH and basal area/ha when plotted against prediction, age, altitude and time increment length (Figures 7.5 and 7.6). A parameter t_0 was added to the position of T_1 and T_2 and it significantly improved the prediction performance for juvenile growth of MTH but it was not included in the model of basal area as it showed no significance. A transformation of the other parameter (β) in the model gave no improvement in fit.

Table 7.6: Model fitting statistics before and after incorporation of altitude

Variable	MSE		Std Deviation		Skewness		Kurtosis	
	Before	After	Before	After	Before	After	Before	After
MTH	0.6968	0.5795	0.83	0.76	0.33	-0.18	2.02	0.83
G	10.3752	6.1461	3.22	2.48	1.25	-0.28	6.90	3.28

A regression analysis of residuals of the final models of MTH and basal area was again conducted against explanatory variables. Table 7.8 shows that none of the explanatory variables had a close correlation with the residuals of MTH and basal area/ha. No additional variables could be included in the model to improve projections significantly. GF rating was marginally significant at the 5% level for basal area/ha and should be studied in the future. The test was valid, as the data were prepared with one random observation in each plot.

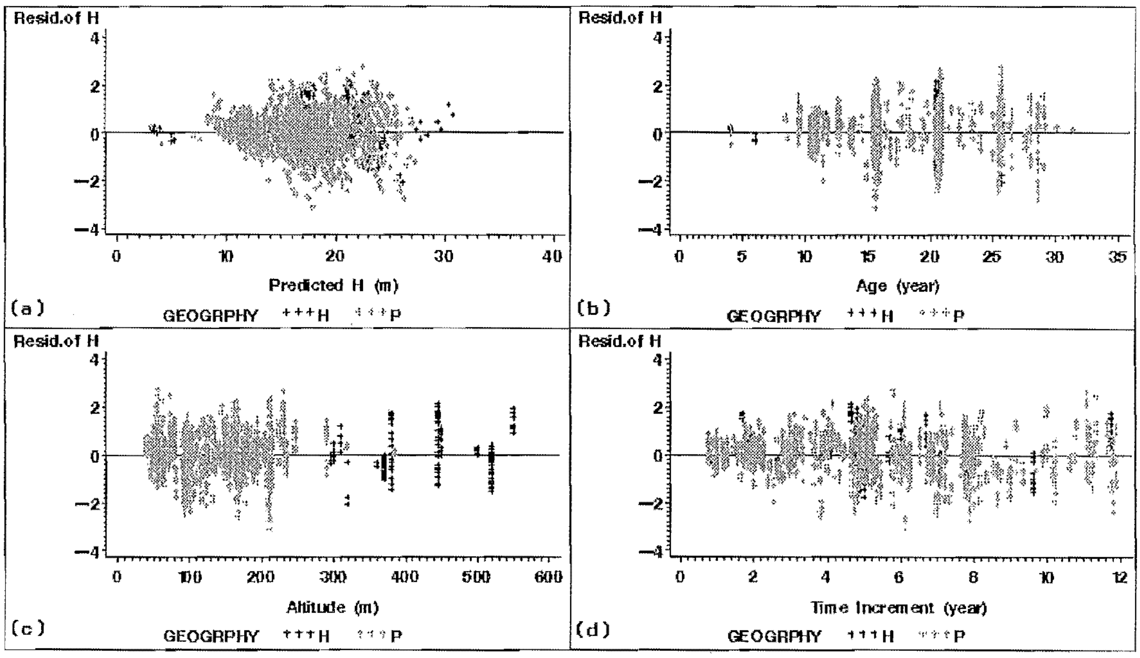


Figure 7.5: The residual pattern of MTH model after inclusion of altitude

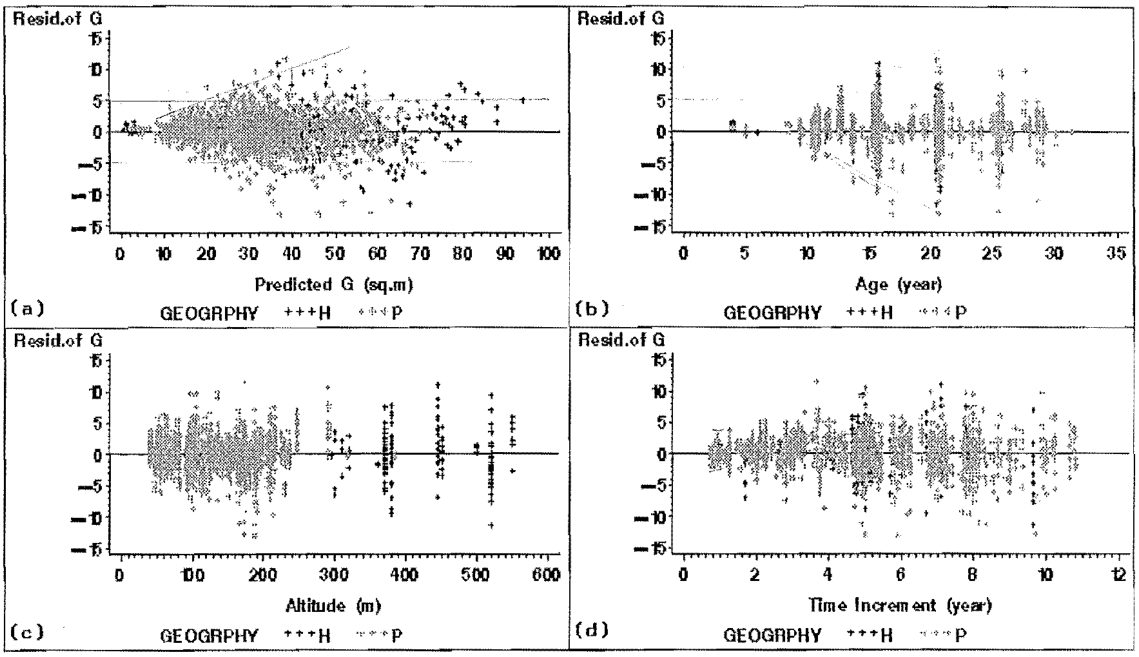


Figure 7.6: The residual pattern of basal area model after inclusion of altitude

Table 7.7: Parameter estimates for equations of MTH and basal area/ha

Equations	Parameter	Estimates	Std. Error	MSE
$H_2 = \exp[\ln(H_1)((T_1+t_0)/(T_2+t_0))^{\beta}$	α_0	4113.5576	60.5857	0.5795
$+(\alpha_0+\alpha_1 Alt+\alpha_2(Alt-250)X)/1000$	α_1	0.501580	0.07200	
$(1-((T_1+t_0)/(T_2+t_0))^{\beta})]$	α_2	0.698224	0.13505	
(X=0 when altitude<250	β	0.718274	0.04761	
X=1 when altitude≥250)	t_0	1.676672	0.40048	
$G_2 = \exp[\ln(G_1)(T_1/T_2)^{\beta}+(\alpha_0+\alpha_1 Alt$	α_0	4520.9455	18.5762	6.1461
$+\alpha_2(Alt-250)X)/1000(1-(T_1/T_2)^{\beta})]$	α_1	0.808093	0.08028	
	α_2	1.450708	0.13657	
	β	1.086840	0.012235	

Table 7.8: Regression test results for residuals of MTH and basal area/ha

Dependent Variable	Independent Variable	DF	Parameter Estimate	Std. Error	T	Prob > |T|
	Intercept	1	-0.01782	0.386673	-0.046	0.9633
	Altitude	1	0.000406	0.000422	0.961	0.3375
Residual	GF rating	1	0.007025	0.020475	0.343	0.7319
of MTH	Pruning height	1	-0.05053	0.072691	-0.695	0.4878
(P>F=0.18)	Initial stocking	1	0.000213	0.000226	0.942	0.3475
	Time increment	1	-0.03187	0.024375	-1.308	0.1925
	Error	197				
	Intercept	1	2.726659	2.756939	0.989	0.324
	Altitude	1	-0.00031	0.003146	-0.099	0.9216
	GF rating	1	0.171917	0.087448	1.966	0.0508
Residual	Pruning height	1	0.322417	0.326501	0.987	0.3247
of G	Site index	1	-0.25143	0.154973	-1.622	0.1064
(P>F=0.23)	Initial stocking	1	0.001684	0.000908	1.854	0.0654
	Time increment	1	-0.04701	0.100322	-0.469	0.6399
	Error	183				

7.3.2 Modelling of stems per hectare

The factors that cause trees to die in Canterbury include drought, windthrow, occasional heavy snow, animals, diseases, and competition from neighbouring trees and weeds. Most of those factors cannot be predicted precisely so surviving trees per hectare is a model that is difficult to fit satisfactorily (Glover and Hool, 1979; Woollons, 1998).

Many difference equations appropriate for projections of stocking have been listed by Clutter *et al.* (1983). To improve prediction, modellers have tried to incorporate environmental variables into the mortality model (Mason 1992, for example) and to use a variant of basic difference equation (Woollons and Hayward 1985).

Six equations were fitted and the results are listed in Table 7.9. Equations 2 and 3 in table 7.9 are varieties of Equation 1. Equation 1 led to the smallest MSE so it was chosen in the study. A regression analysis of residuals derived from Equation 1 was conducted. Table 7.10 showed that none of the explanatory variables had a significantly close relationship with residuals of stems per hectare. Altitude was close to significance at 5%. Its effect on survival was far less than on basal area and MTH. Not much improvement could be obtained with its inclusion in the model. Projections with Equation 1 were not as precise as those of MTH and basal area, as shown with residual patterns in Figure 7.7. Little bias was apparent, however. Ninety percent (>5% and <95%) of residuals were within ±50 stems/ha for projection lengths up to 11.5 years.

Table 7.9: Equations and fitting statistics for stems per hectare

No.	Equations	Parameter	Estimates	Std. Error	MSE
1	$N_2=(N_1^c + a/100000\ (T_2^b - T_1^b))^{(1/c)}$	a	0.0437028	0.01856	1102.6
		b	1.2936403	0.09958	
		c	-1.304660	0.06852	
2	$N_2=(1/N_1^{0.5}+a/100000(T_2^b-T_1^b))^{(-2)}$	a	0.5295299	0.17723	1156.2
		b	1.8923010	0.09772	
3	$N_2=(1/N_1^{0.5}+a((T_2/100)^2-(T_1/100)^2))^{(-2)}$	a	0.03666865	0.00088	1156.4
4	$N_2=N_1\ exp\ [\ b(T_2-T_1)]$	b	-0.00660982	0.00018	1311.4
5	$N_2=N_1\ (T_2/T_1)^a\ exp\ [b\ (T_2-T_1)]$	a	0.12172703	0.00965	1236.0
		b	-0.01496154	0.00069	
6	$N_2=N_1\ exp\ (a\ (T_2^b-T_1^b))$	a	-0.0001042	0.000037	1235.0
		b	2.21242348	0.104643	

Total df = 2601

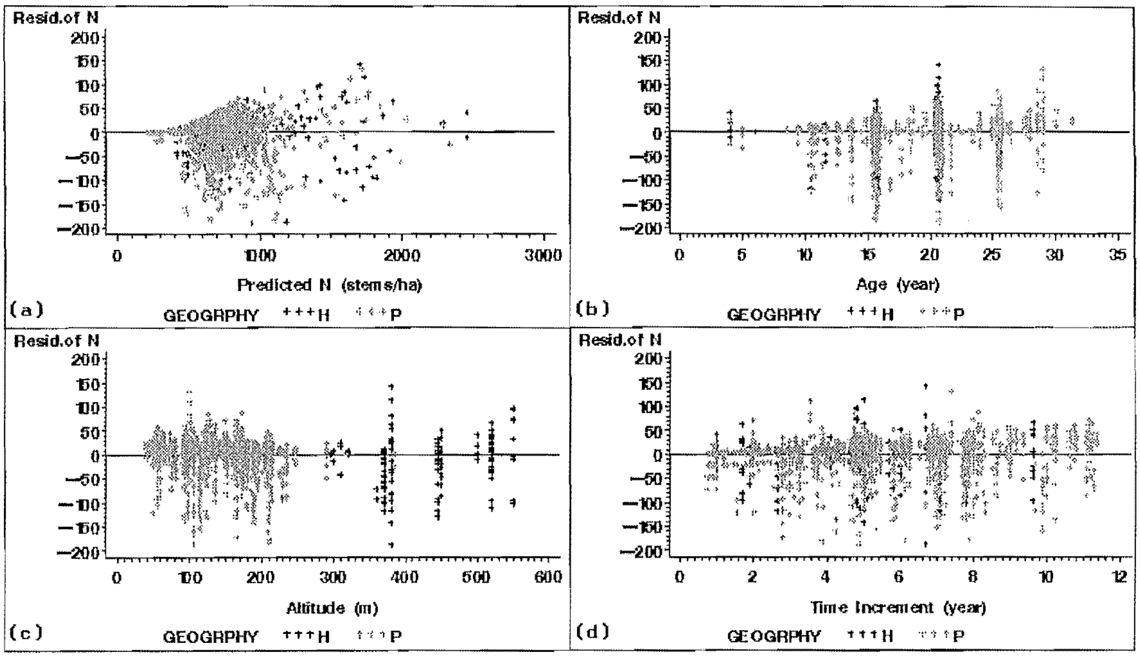


Figure 7.7: The fit of projection equation of stems per hectare

Table 7.10: Regression test results for residuals of stems per hectare

Dependent Variable	Independent Variable	DF	Parameter Estimate	Std. Error	T	Prob > |T|
	Intercept	1	-39.7303	33.8029	-1.175	0.2413
Residual	Altitude	1	-0.0541	0.0276	-1.962	0.0511
of N	GF rating	1	0.9656	1.0517	0.9180	0.3597
(R^2=0.04)	Pruning height	1	0.3918	3.6241	0.1080	0.9140
(Prob>F=0.16)	Site index	1	1.6186	1.9587	0.8260	0.4096
	Time increment	1	1.5645	1.3062	1.1980	0.2325
	Error	194				

7.3.3 Modelling of volume per hectare

A model of volume per hectare was built using data from SPBL. Tree-volume equations are usually accurate only for trees far taller than 1.40 m, so data for juvenile growth were not included when modelling volume. No volume information was available for plot measurements in Eyrewell forest.

Stand volume per hectare is usually estimated by its strongly related state variables such as basal area, mean top height or site index, age, and stocking at any age (Clutter, *et al.*, 1983). The most used stand volume equation in New Zealand is the linear model of the ratio of volume to basal area against MTH (i.e. $V/G=\alpha+\beta H$ in McEwen, 1978). It did not perform so well for this data set, however, and a bias appeared (Figure 7.8).

An alternative equation form (Equation 2 in Table 7.11) that was used by others as a tree volume equation (Brackett, 1973) and stand volume equation (Temu, 1992), displayed a better fit for this study. A reduction of 30% of mean square error was gained compared with the traditional model (Equation 1 in the Table 7.11). Figure 7.9 presents the improved residual patterns for Equation 2 in Table 7.11. Table 7.12 lists the results of residual regression analysis. No explanatory variable showed a close correlation with the volume residuals.

Future volume could be estimated using this model and the established projection equations of basal area and MTH. Figure 7.10 shows the fit of volume projections. Up to an 11 years projection period was included, so the range of residuals was much higher than that in Figure 7.9. The most important finding was that no obvious bias appeared with the projection.

Table 7.11: Equations and fitting statistics for volume per hectare

Equations	Parameter Estimates & Std. Error			MSE	Skewness
(1) $V=G(\alpha+\beta H)$	α	1.082447	0.006416	16.0406	-3.8646
$(V/G=\alpha+\beta H)$	β	0.293274	0.000433		
(2) $V=\alpha G^{\beta} H^{\gamma}$	α	0.6225102886	0.002992	11.2761	-1.00318
(where $H=MTH$)	β	0.9670398052	0.001411		
	γ	0.8466802294	0.002534		

Table 7.12: Regression test results for residuals of volume per hectare

Dependent Variable	Independent Variable	DF	Parameter Estimate	Std. Error	T	Prob > |T|
Residual of V (R^2=0.0075) (Prob>F=0.96)	Intercept	1	-0.45499	3.73175	-0.122	0.9031
	Age	1	-0.00417	0.06556	-0.064	0.9494
	Altitude	1	-0.00099	0.00241	-0.413	0.6800
	GF rating	1	0.05831	0.10419	0.560	0.5764
	Pruning height	1	0.30408	0.35066	0.867	0.3869
	Initial Stock	1	0.00074	0.00108	0.688	0.4925
	Site index	1	-0.03994	0.18572	-0.215	0.8299
	Error	193				

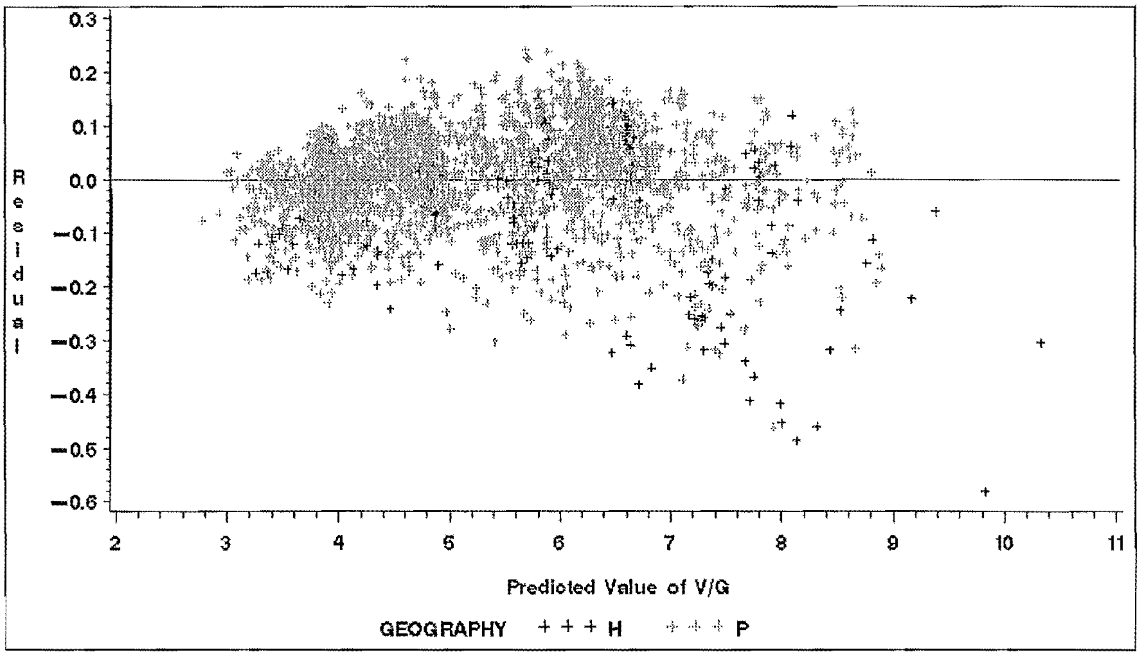


Figure 7.8: The fit of volume equation of $V / G = \alpha + \beta H$

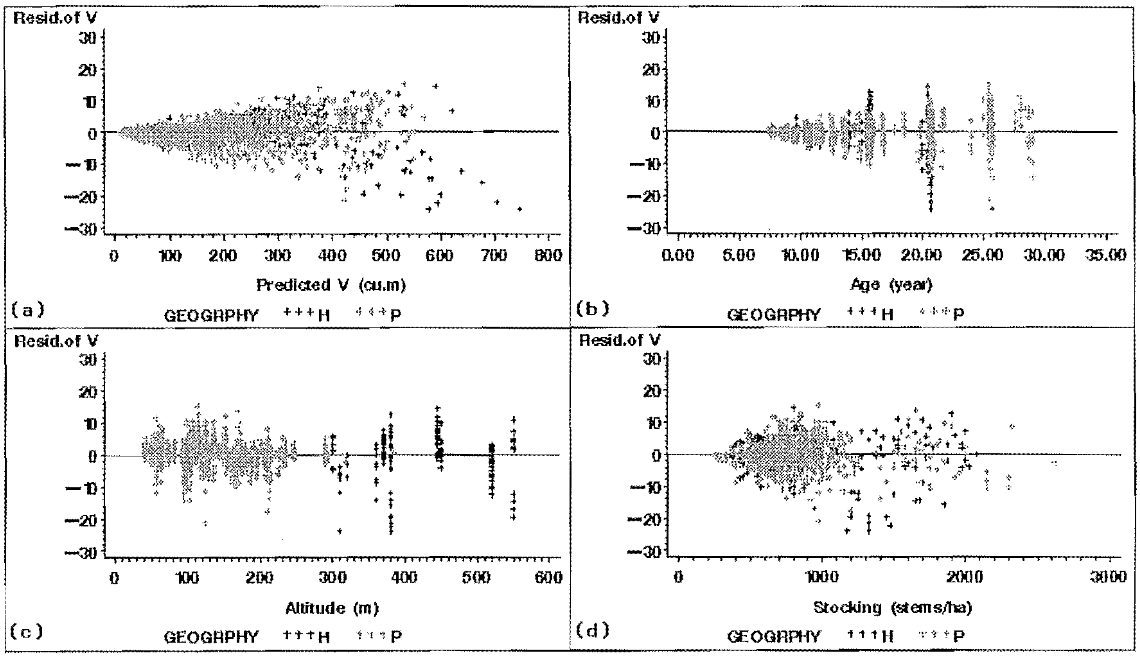


Figure 7.9: The fit of volume equation of $V=\alpha\ G^{\beta}H^{\gamma}$

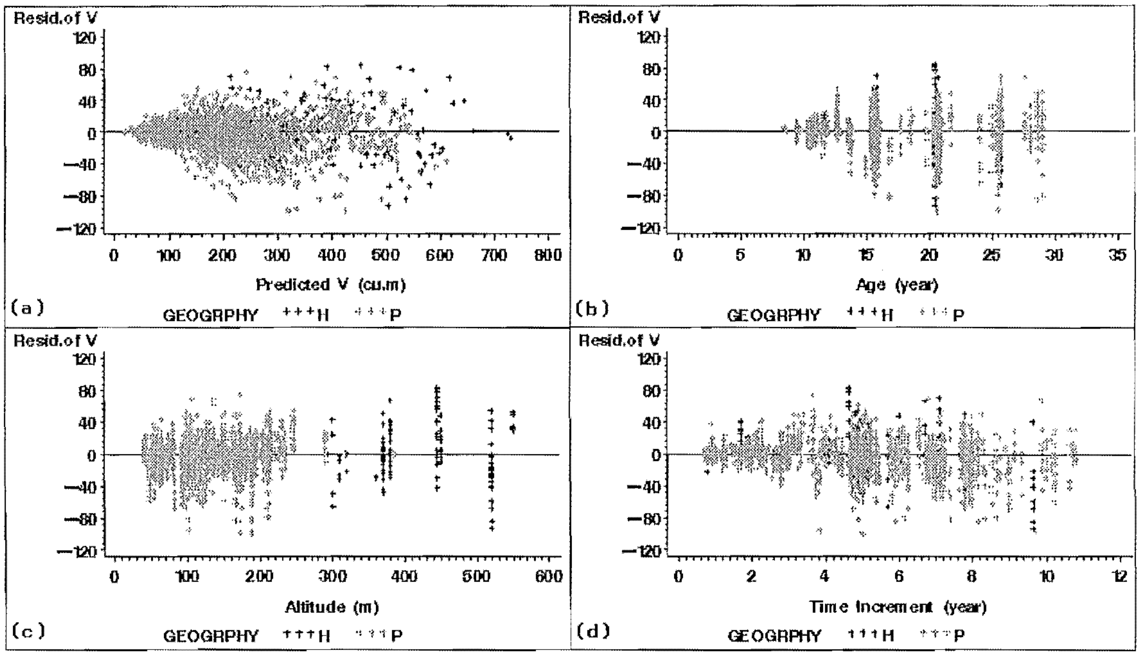


Figure 7.10: The fit of volume projections with up to 11 years' time increments

7.3.4 Projection of stand tables with diameter distributions

This study employed the method of moments in combination with the parameter recovery method to estimate and project the three parameters for a reverse Weibull distribution at a stand level. Projection equations for standard deviation and maximum diameter were the elements needed to structure the stand-table projection system completely. The maximum diameter in a stand was selected from all sampled trees in the stand. Standard deviation of diameters in a stand was estimated with a cluster sampling formulation of Equation 3.3 in Chapter 3, which accounts for the variations within and between plots.

Only data from SPBL's estate were used for this portion of the study since the concept of stand and plot was well defined in SPBL's database. Stands with the number of plots less than three were dropped because of the need of a reliable estimate for standard deviation and maximum diameter.

7.3.4.1 An analysis of stand and plot variance

It was pointed out by Garcia (1991) that stand variance could be underestimated with the average of plot variances in a stand and that it is theoretically correct to use the cluster sampling formulation. The magnitude of bias was usually not serious in practice but it needed to be investigated (Whyte and Woollons, 1992).

The material in this study's database provided an opportunity to resolve the issue. The means of plot variance and stand variance estimated with a cluster sampling formulation were all obtained. Table 7.13 lists the summary from 402 stand measurements of the mean of plot variance, stand variance, the difference between plot and stand variance and relative bias as a percentage. Average bias was -0.93 cm^2 and relative bias ranged from -50.2% to 12.4%, with an average of -8%. Ten percent of stand measurements were seriously underestimated, with a relative bias between -50.2% and -21%. Figure 7.11 shows the differences between the mean of plot variance and stand variance. In most cases the means of plot variance were lower than stand variances. Bias was not serious for most stands but was serious for a few stands.

Table 7.13: Summary of average plot variance, stand variance and bias

Variable	Frequency	Mean	Std.dev.	Min.	Max.
Mean plot variance (1)	402	14.82	12.53	1.57	65.14
Stand variance (2)	402	15.75	12.65	1.62	67.13
Bias (1) - (2)	402	-0.93	1.44	-7.37	4.28
Relative bias [(1)-(2)]/(2)	402	-8.0%	10.1%	-50.2%	12.4%

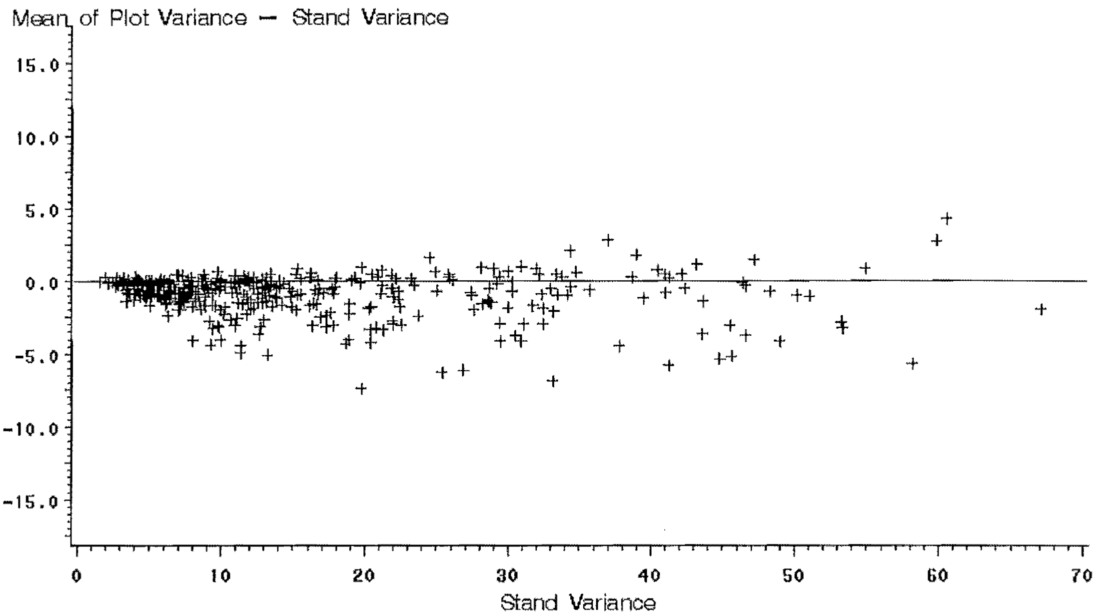


Figure 7.11: The difference between the mean of plot variance and stand variance

7.3.4.2 Projections for standard deviation and maximum diameter

After fitting equations listed in Table 7.1, a polymorphic Schumacher equation was chosen as a basic form for both maximum diameter and standard deviation. Altitude entered the equations with a linear form. Table 7.14 shows the final equation forms, parameter estimates and fitting statistics. Figure 7.12 and 7.13 show residual patterns and little bias was apparent. For projection lengths up to 10 years, 98% (>1% and

<99%) of residuals for maximum diameter projections were within ±4.5 cm while 98% of residuals for standard deviation projections were within ±0.85 cm.

Table 7.14: Equations and fitting statistics for maximum dbh and standard deviation

Equation		Parameter Estimates	Std. Error	MSE	Error Skewness
$d_{\max 2} = \exp[\ln(d_{\max 1})(T_1/T_2)^{\beta} + (\alpha_0 + \alpha_1 Alt)/1000\ (1-(T_1/T_2)^{\beta})]$	α_0	4223.7965	47.5049	2.4280	0.47
	α_1	0.791615	0.1092		
	β	0.86580	0.04414		
$d_{std 2} = \exp[\ln(d_{std 1})(T_1/T_2)^{\beta} + (\alpha_0 + \alpha_1 Alt)/1000\ (1-(T_1/T_2)^{\beta})]$	α_0	3475.6691	236.2146	0.1002	-0.07
	α_1	1.275058	0.29085		
	β	0.374693	0.03819		

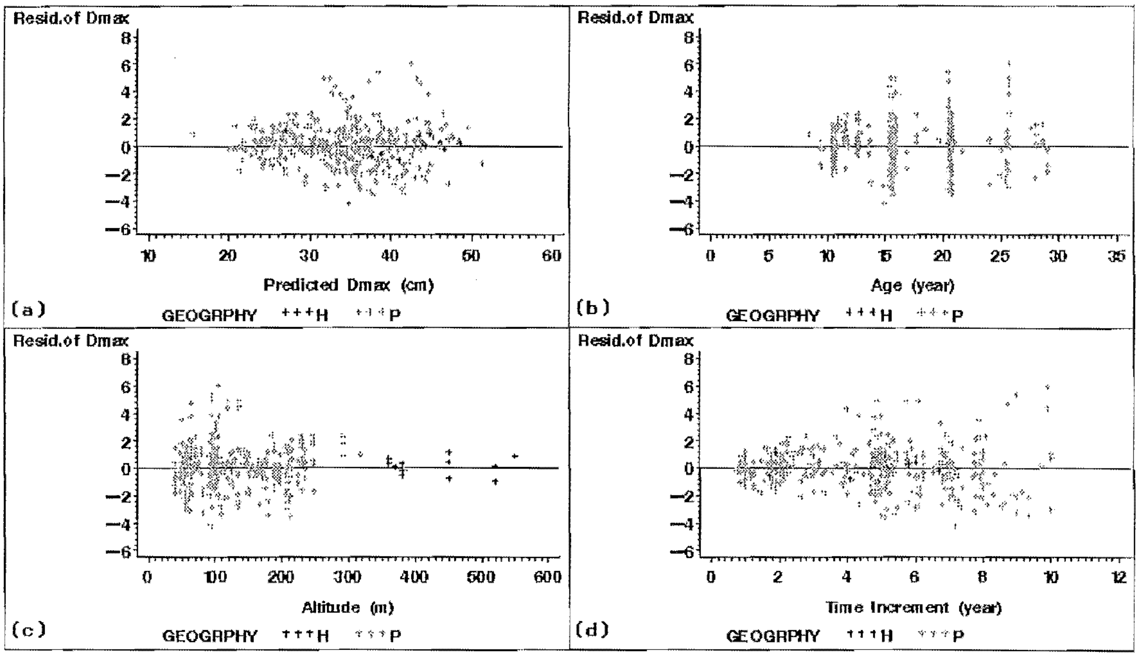


Figure 7.12: The fit of maximum diameter projection equation

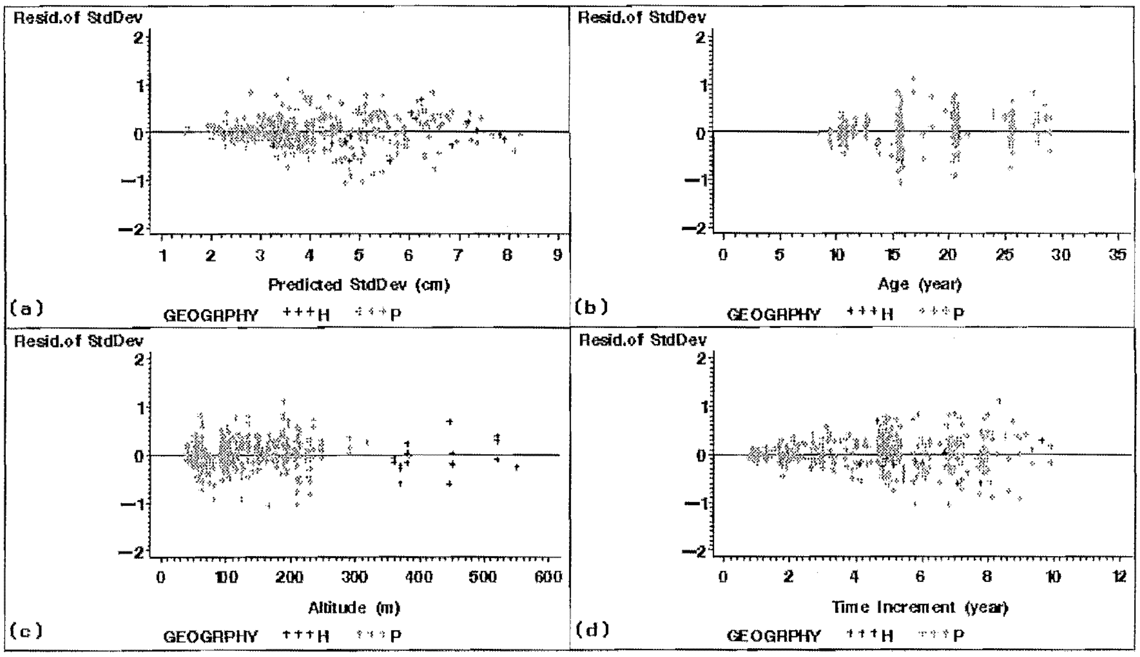


Figure 7.13: The fit of standard deviation projection equation

7.3.4.3 The fit of dbh distribution projection and the construction of stand tables

At this point, all model components were ready for the projection of dbhob distributions. Table 7.15 lists the necessary equation forms to project stand tables. The three parameters in the reverse Weibull equation had been calculated for all individual stand measurements. The location parameter "*a*" could be estimated by a maximum diameter projection equation. The scale parameter "*b*" and shape parameter "*c*" could be derived from Garcia's (1981) method using a standard deviation projection equation and the projection equation of basal area per hectare and living stems per hectare that had been built up previously. The predicted values were calculated with a Weibull cumulative density equation.

Diameters were sorted and counted by 1-cm classes for each stand measurement and the percent and cumulative percentage for each class were then calculated as the actual values. Projected values were compared with the actual values and the fit of the cumulative diameter distribution model is illustrated in Figure 7.14. Little bias was apparent.

Table 7.15: Equations necessary for projection of stand tables

Number and name		Equation form
1	Weibull density	$f(D) = \frac{c}{b}[(\frac{a-D}{b})^{c-1} \exp[-(\frac{a-D}{b})^{c}]$ (D = dbh class, a, b, c = parameter)
2	Weib.-cumulative	$F(D) = \exp[-(\frac{a-D}{b})^{c}]$ (D = dbh class, a, b, c = parameter)
3	Std. deviation of stand	$d_{std} = \sqrt{[\frac{1}{m\bar{n}}\Sigma\Sigma d_{pi}^{2} + \frac{1-m/M}{m\bar{n}(m-1)} n_p \bar{d}_p^{2} - \frac{1-1/M}{m\bar{n}^{2}(m-1)b}(\Sigma n_p \bar{d}_p)^{2}]/(1-1/\bar{n}M)}$ (d_{pi}=i^{th} dbh in p^{th} plot, m=sample size, M=potential plots, n_p= the number of trees in p^{th} plot, $\bar{n}$ =average of number of trees)
4	Arithmetic mean	$\bar{d} = \sqrt{d_{\bar{g}}^{2} - d_{std}^{2}} = \sqrt{\frac{40000G}{\pi N} - d_{std}^{2}}$ ($d_{\bar{g}}$=diameter of mean basal area, G=basal area/ha,N=stems/ha)
5	Maximum Dbh	$d_{\max} = \max(d_{pi})$ (d_{pi} = i^{th} dbh in p^{th} plot)
6	Location a	$a = d_{\max}$
7	Scale b	$b = [\Gamma(1+1/c)/(a-\bar{d})]^{-1}$
8	Shape c	$c = \{z[1+(1-z)^{2}(-0.2200991 - 0.00194664\ z + 0.15310925\ z^{2} - 0.08354348\ z^{3} + 0.007454537\ z^{5})]\}^{-1}, z = d_{std}/(a-\bar{d})$
9	Projection of Std. deviation of dbh	$d_{std\,2} = \exp[\ln(d_{std\,1})(T_1/T_2)^{\beta} + (\alpha_0 + \alpha_1 Alt)(1-(T_1/T_2)^{\beta})]$ (T_1,T_2= initial and projection age, Alt = altitude)
10	D_{max} projection	$d_{\max 2} = \exp[\ln(d_{\max 1})(T_1/T_2)^{\beta} + (\alpha_0 + \alpha_1 Alt)(1-(T_1/T_2)^{\beta})]$
11	Height curve	$h = 1.40 + [0.695955 + 0.666983T^{-0.5} - 0.106771\ln(SI) + (0.954201 + 0.000741Alt)/d]^{-5}$ (h = height, T = age, SI = site index, Alt = altitude, d = diameter)

To project an output of stand table with the columns of diameter class, frequency and class mean height, the height curve established in chapter 4 (Equation 11 in Table 7.15) can be used. Height for each diameter class is a function of the diameter class, stand age, site index and altitude. The number of stems for each class is a product of total stems per hectare and predicted proportion for the class. Volume can be computed for a known diameter class and class height using the tree volume equation that was introduced in chapter 3.

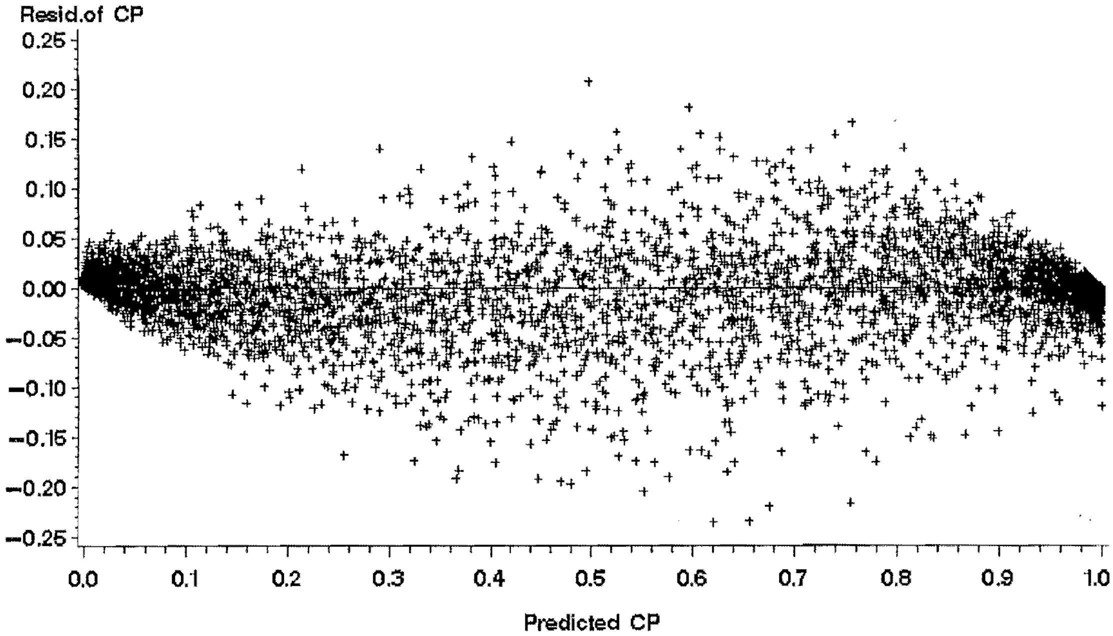


Figure 7.14: The fit of projected diameter distribution for all individual stands

(CP = cumulative probability)

7.3.5 Examination of main model components

Two sources of data at a stand and plot level were used to examine the performance of model components MTH, basal area per hectare, stems per hectare, and volume per hectare. One was a set of independent plot measurements chosen randomly before building the new model as was described in chapter 3. In total there were 293 observations in the data set, which was independent of the data used in building the model. The second source of the data was the same as data used in the model establishment but prepared at stand-level. Not all plots could be included to derive the statistics for a stand. To keep a stand with consistently the same plots in all historical re-measurements, some newly established plot measurements or plot measurements with one or two measurements missing had to be dropped. A total of 225 observations was included in the data set.

7.3.5.1 Model examination at a plot level

Table 7.16 shows the model performance measures of average model bias, effective factor and skewness of residuals for the new established model. All values of EF were high and all values of AMB were quite low compared with predictions for the four variables of mean top height, basal area, stems per hectare and volume. Skewness of residuals was reasonably low.

Figure 7.15 shows the residual pattern against predictions from the new model. No apparent bias could be detected for any of the four model components. Projections for stems per hectare, a component that is difficult to fit satisfactorily, were less precise.

Table 7.17 shows the results of regression of residuals with a set of autocorrelation-free data. 58 observations out of a total of 293 were included in the data set. The results indicated that none of the explanatory variables was significantly related to residuals of MTH, basal area, stems per hectare, and volume. The correlation between residuals and predictions was not significant (see last two rows in Table 7.17).

Table 7.16: The results of model performance measures

Model component	Plot level			Stand level		
	AMB	EF	Skewness	AMB	EF	Skewness
MTH	0.292	0.970	-0.134	0.114	0.968	-0.250
G	-0.190	0.975	0.344	0.392	0.988	0.103
N	3.116	0.979	-1.56	0.159	0.987	-0.783
V	0.1565	0.988	0.292	1.869	0.994	-0.031

Table 7.17: Regression test results with plot measurements

Variable	H		G		N		V	
	Coefficient	Pr>lTl	Coefficient	Pr>lTl	Coefficient	Pr>lTl	Coefficient	Pr>lTl
Intercept	-0.0568	0.975	8.7878	0.1003	56.2901	0.3523	30.0518	0.4083
Age	0.0238	0.5045	0.0600	0.5697	-0.3855	0.7489	0.7203	0.3218
Time Inc	0.0582	0.5468	-0.1237	0.6653	1.1774	0.7184	-1.3484	0.4928
Altitude	0.0000	0.9924	0.0009	0.7839	-0.0414	0.265	-0.0127	0.5694
GF rating	0.0488	0.1149	0.1121	0.2197	0.2319	0.8229	0.7005	0.2636
$H_{pruning}$	0.3460	0.0981	-0.0203	0.9735	1.2926	0.8533	1.1788	0.7792
Site Index	-0.0803	0.3651	-0.5025	0.0594	-2.4335	0.4174	-2.1118	0.2435
Intercept	-0.3295	0.5845	-0.0551	0.9514	-10.2297	0.4678	-1.5767	0.749
Prediction	0.0257	0.4221	0.0009	0.9683	0.0203	0.2845	0.0063	0.7062

Where Time Inc is time increment (projection length), $H_{pruning}$ is pruning height

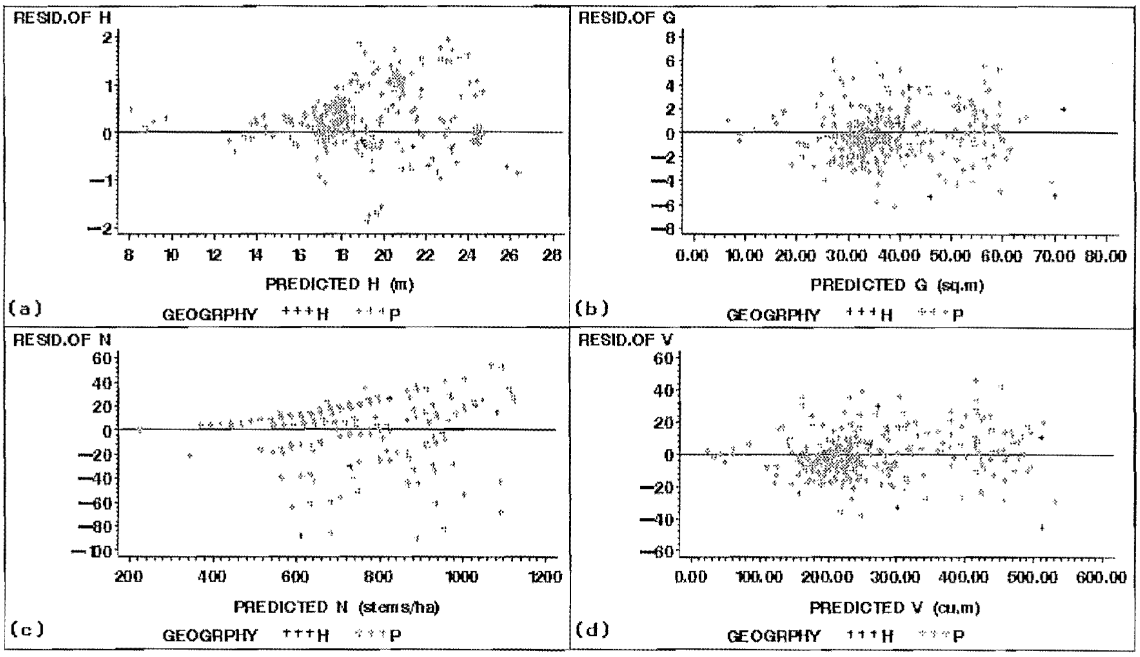


Figure 7.15: The fit of the new model to plot data set

7.3.5.2 Model examination at a stand level

Figure 7.16 shows the residual pattern of the new model for all stand measurements. The model predicted the data reasonably well. No large bias appeared. Table 7.18 shows the results of sample size weighted regressions of residuals with a set of autocorrelation-free data. 129 observations out of a total of 225 were included in the data set. The results indicated that the correlation between residuals and predictions was not significant (seen in the last two rows of Table 7.18). None of the explanatory variables was significantly related to residuals of basal area and volume but projection length and altitude were significantly related to residuals of MTH, and altitude and pruning height were significantly related to residuals of stems per hectare at the 5% level. The trends of the significant correlation were not so serious, however, as can be seen with the R-square value (the first row in Table 7.18) for the overall regressions. The trends were checked and inclusion of those variables into the model could not be justified. They should be further studied in future, however.

Table 7.18: Regression test results with stand measurements

Variable	H (R^2=0.10)		G (R^2=0.08)		N (R^2=0.11)		V (R^2=0.07)	
	Coefficient	Pr>|T|	Coefficient	Pr>|T|	Coefficient	Pr>|T|	Coefficient	Pr>|T|
Intercept	0.0477	0.9639	2.5262	0.3060	22.636	0.6314	12.716	0.4237
Age	0.0306	0.1205	0.0536	0.2439	0.4237	0.6298	0.3908	0.1878
Time Inc	-0.1102	0.0097	-0.1588	0.1078	1.2987	0.4903	-1.0641	0.0947
Altitude	0.0013	0.0442	0.0025	0.1093	-0.0874	0.0038	-0.0027	0.7839
GF rating	-0.0010	0.9496	0.0262	0.5007	0.9534	0.2015	0.1122	0.6543
$H_{pruning}$	0.0623	0.4445	-0.0799	0.6743	-8.2854	0.0243	1.9230	0.1184
Site Index	-0.0131	0.7883	-0.1371	0.2301	-0.4127	0.8500	-0.9160	0.2138
Intercept	0.1099	0.6706	-0.1187	0.7126	-12.977	0.1559	1.1699	0.4834
Prediction	0.0038	0.8051	0.0150	0.0965	0.0182	0.1257	0.0015	0.8172

Where Time Inc is time increment (projection length), $H_{pruning}$ is pruning height

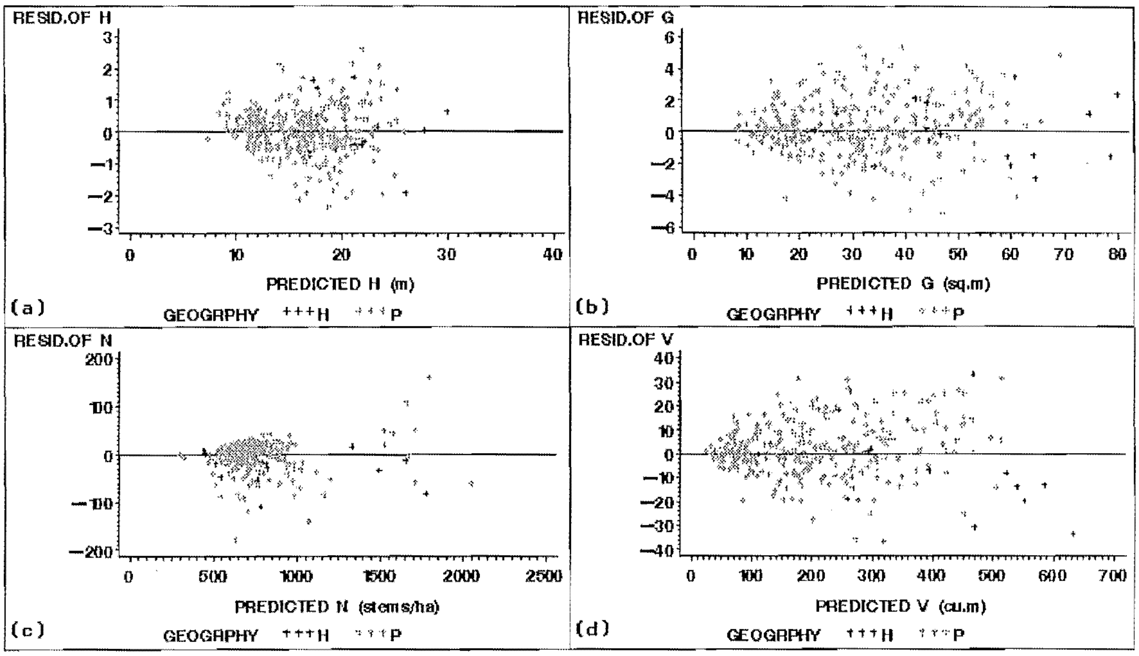


Figure 7.16: The fit of the new model to stand level data set

In summary, the residual analyses showed that the equations for MTH, basal area per hectare, stems per hectare and volume per hectare showed little apparent bias for use at a plot level and a stand level. Early studies by Hann (1980), Moeur and Ek (1981), and Smith and Burkhart (1984) found that yields estimated from projections of individual plots were more accurate than projections from the means of all these plots. For uniform, even-aged stands, however, Hägglund (1981) observed that projections of stand means provided the same estimates as averaged projections of each individual plot, which was in agreement with this study.

7.4 Conclusions

The stand model components of mean top height, basal area/ha, stems/ha, volume/ha and diameter distribution were separately developed by selecting suitable equations, using appropriate data structures, and identifying and incorporating significant explanatory variables to improve prediction precision. A polymorphic Schumacher equation displayed the best fit for both MTH and basal area. A data structure of all possible intervals was chosen to fit all difference equations because it was a structure that resulted in long projections of MTH with least bias. The altitude effect was formulated and included in the selected equation and it significantly improved the model with a reduction of mean square error by 17% for MTH and 41% for basal area. A three-parameter difference equation was chosen for stems per hectare. Altitude was close to but less than the significant level of 5%. Its effect on survival was far less than on basal area and MTH. Not much improvement could be obtained with its inclusion in the model. The stand volume equation form of $V=\alpha G^{\beta}H^{\gamma}$ displayed a better fit than the traditional form of $V/G=\alpha+\beta H$ and a reduction of mean square of error by 30% was gained. Future volumes were estimated with the projected future basal area/ha and future MTH. As a result no apparent bias could be detected.

Projection equations of maximum diameter and standard deviation were built and a system for producing stand tables with a reverse Weibull distribution was established. A projection of diameter distribution for all individual stands was conducted and it resulted in an unbiased fit. An analysis of underestimation bias for stand variance using the mean of plot variance showed that the bias was not serious for most stands but was for a few stands. On average the difference between the mean of plot variance and stand variance was -0.93 cm^2.

Two sources of data at a plot level and a stand level were used to examine the performance of the model. The residual analysis showed that the model gave no significant bias for MTH, basal area per hectare, stems per hectare and volume per hectare in use at both a plot level and a stand level in SPBL's estate.

CHAPTER 8

DEVELOPMENT OF MODEL CanSPBL (2): TREE MODEL

8.1 INTRODUCTION

Tree models use the individual tree as the basic modelling unit. New developments in the representation of wood qualities are causing managers to demand more precise models of the growth and yield of individual trees. Recent advances in computer technologies have meant lower costs of building and using tree models. Tree models can depict irregular non-unimodal stands much more realistically than a Weibull distribution function can.

Munro (1973) identified two classes of individual-tree models: distance-dependent individual tree models and distance-independent individual tree models. Whether or not tree spatial information improves predictions of individual-tree growth has been much debated in the literature (Daniels and Burkhart, 1975; Clutter *et al.*, 1983; Bruce and Wensel, 1987; Biging and Dobbertin, 1995). Any advantage of distance-dependent individual-tree models depends on the reliability of competition measures that might be of lesser value for plantations in which full control of spacing is exercised. For this study the unavailability of inter-tree distances limited the application of such models.

A tree can either grow to a certain size or die in a given period, so tree models may include the components of tree diameter (basal area or relative basal area as its alternative), height, volume, and mortality. Most individual-tree models describe the increment of diameter, but a few models project diameter and height directly based on difference equations (Zhang *et al.*, 1996a, for example). This chapter focuses on the modelling of dbhob and the survival of individual trees. Height can be estimated with diameter, age, site index and altitude using the regional height-diameter equation that was reported in Chapter 4. Tree volume can be estimated with dbh and height using a tree-volume equation described in Chapter 3.

For diameter modelling, an approach using relative basal area projection from an initial individual-tree list or stand table (Clutter and Jones, 1980; Pienaar and Harrison, 1988) had been demonstrated as a superior method compared to Weibull distribution and percentile-based methods in projecting stand tables (Borders and Patterson, 1990). Tree diameters without disturbance might follow a sigmoid path, so modelling with sigmoid difference equations might provide a suitable approach as well.

For mortality modelling, logistic regression is most commonly employed in individual-tree model systems to estimate survival probability of each tree (Hamilton and Edwards, 1976; Hann, 1980; Hamilton, 1990; Vanclay, 1991; Avila and Burkhart, 1992; Zhang *et al.*, 1997). The projected value with logistic equation is bounded between 0 and 1. The inclusion of good predictors into models could increase prediction reliability.

An individual-tree model should not be separated from a stand level model but be regarded as complementary option (Burkhart, 1977; Daniels and Burkhart, 1988). Somers and Nepal (1994) proposed a link between an individual-tree model and a stand model based on the assumption that stand level components were estimated correctly by the stand level model and that individual-tree growth should be adjusted. Stand models have been generally found to be precise and unbiased in New Zealand (Garcia, 1991; Goulding, 1994). This study adopted the same view and disaggregative adjustments were made for tree size projections when the sum of individual trees in a plot was not the same as projections based on a projection equation at stand level (as described in the previous chapter).

Finally in this chapter, a computerised simulation program for use at both tree level and stand level was developed to provide a tool to obtain quantitative descriptions of future production. It yields projections for multiple stands or plots simultaneously given recent measurements.

8.2 Data for Modelling Tree DBH and Survival

The data used were those plots of 0.04 ha in area and where at least 10 trees remained in each plot. A maximum of 5 plots per stand was included to keep the number of sampled trees not too excessive. All trees were uniquely identified. Table 8.1 lists a summary of the data set for a total of 47 731 trees.

The data was reorganised to fit projection equations (difference equations). A data structure of all possible intervals (time increments) consisting of 45 456 observations was used. The minimum interval length was 0.8 year and the maximum was 10.5 years.

Table 8.1: Summary of data for individual-tree modelling

Variable	Frequency	Mean	Std. dev	Minimum	Maximum
Age	47 731	14.41	5.16	7.45	29
dbhob (d)	47 731	20.6	6.4	2.3	50.9
Relative diameter (Rd)	47 731	1.00	0.17	0.10	1.92
Relative basal area (R)	47 731	1.00	0.33	0.02	3.29
Pruning height	47 731	2.5	0.7	0.0	6.0
Altitude	47 731	148	89	40	520
Actual basal area (G/ha)	47 731	28.5	16.1	3.0	89.4
Actual stocking (N: stems/ha)	47 731	778	248	275	2 075

8.3 Methods of Fitting DBH and Survival

For diameter projection, two types of equations were fitted and compared. The first was based on relative basal area, and the other was based on commonly used sigmoid difference equations that had been used for stand models. The same fitting methods were applied to individual-tree diameter as those for MTH and basal area in the stand model.

Borders and Patterson (1990) employed a relative basal area equation (Equation 8.1) and demonstrated the equation performed well for projecting stand tables. A similar form of equation was proposed here in this study and it can be written as Equation 8.2. Change in relative basal area will be dependent on change in T_1 and T_2. R_2 will equal to R_1 when T_2 equals T_1. The two equations have the property of path-invariance. Diameters can be derived and they can be written as Equation 8.3 and 8.4, respectively.

$$R_2 = R_1^{(T_2/T_1)^{\beta}} \tag{8.1}$$

$$R_2 = R_1 \exp[\alpha(\tfrac{1}{T_1} - \tfrac{1}{T_2})] \tag{8.2}$$

$$d_2 = \sqrt{\tfrac{1}{0.00007854}\tfrac{G_2}{N_2}[0.00007854 d_1^2 /(\tfrac{G_1}{N_1})]^{(T_2/T_1)^{\beta}}} \tag{8.3}$$

$$d_2 = d_1\sqrt{\tfrac{G_2 N_1}{G_1 N_2}\exp[\alpha(\tfrac{1}{T_1} - \tfrac{1}{T_2})]} \tag{8.4}$$

where R_1 and R_2 are initial and future relative basal area of trees, T_1 and T_2 are initial and future age, d_1 and d_2 are initial and future tree diameter at breast height out of bark (cm), G_1 and G_2 are initial and future basal area (m^2) at a plot level, N_1 and N_2 are initial and future live stocking at plot level, and α and β are parameters.

Actual G_2 and N_2 at a plot level were to be used to fit diameter equations to isolate bias that may be introduced from projections using equations for plot level. Trees alive at both initial stage (T_1) and projection stage (T_2) were included in the modelling process, therefore, the initial N_1 here should be the same as N_2. Predicted G_2 and N_2 using projection equations were to be entered to compare model overall performance with sigmoid difference equations, however.

For the mortality of individual trees, a logistic regression procedure that employed maximum likelihood estimates was employed. Trees alive at an initial age (T_1) but which were dead at the projection age (T_2) were identified. It was hypothesised that tree death was correlated to age, altitude, initial stocking, projection length, site index, GF rating, relative diameter or relative basal area, and the interaction between them.

8.4 RESULTS FOR FITTING DBH AND SURVIVAL

8.4.1 Modelling of diameter at breast height

Table 8.2 shows the fitted MSE for models based on relative basal area and Table 8.3 shows the fitted MSE for models based on sigmoid difference equations.

For relative basal area based equations, parameter estimates and MSE were given in Table 8.2. The first listed MSE was obtained with actual basal area G_2 and stocking N_2 at plot level, the other was obtained with predicted G_2 and N_2 using plot level equations established in the last chapter. An increase in MSE of 16% resulted from the inclusion of projection errors at plot level. The listed parameter estimates reflect the fitting results using actual basal area G_2 and actual stocking N_2. Two employed equations fitted equally well though Equation 8.3 led to slightly lesser MSE than Equation 8.4.

Graphical plots in Figure 8.1 are shown for Equation 8.3, in which predicted G_2 and N_2 from equations for plot level were used. Residual patterns were well distributed against prediction, age, altitude and projection length. For up to 10.5 years' projection, 98% of residuals were within ±5.0 cm and 90% within ±3.1 cm.

Table 8.2: Fitting statistics for dbh equations based on relative basal area

No.	Equation		Parameter Estimates	Parameter Std.error	MSE (1)	MSE (2)
8.3	$d_2 = \sqrt{\frac{1}{0.00007854}\frac{G_2}{N_2}[0.00007854 d_1^2 / (\frac{G_1}{N_1})]^{(T_2/T_1)^\beta}}$	β	-0.15978	0.0064	3.228	3.862
8.4	$d_2 = d_1 \sqrt{\frac{G_2 N_1}{G_1 N_2}\exp[\alpha(\frac{1}{T_1} - \frac{1}{T_2})]}$	α	-0.19810	0.0223	3.272	3.891

(1) Obtained with actual G_2 and N_2, (2) obtained with predicted G_2 and N_2

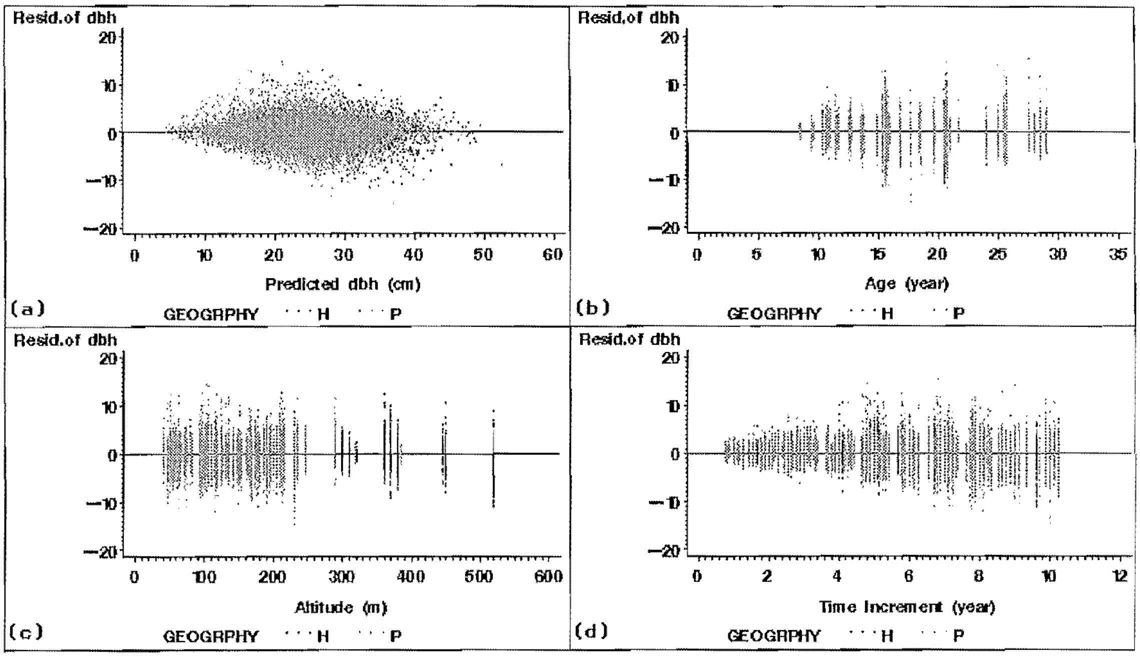


Figure 8.1: The fit of dbh projections for a model based on relative basal area

Among various sigmoid difference equations listed in the last chapter for modelling mean top height and basal area/ha, the anamorphic Schumacher equation and anamorphic monomolecular equation led to the least MSE for tree diameter (Table 8.3). When the anamorphic Schumacher equation was expanded into Equation 8.18 with the inclusion of significant explanatory variables, the MSE of the model was 4.4 which is a reduction of 12%. Graphical plots for the fit of the equation are shown in Figure 8.2. Ninety-eight percent of residuals were within ±5.7 cm and 90% of within ±3.5 cm for up to 10.5 years of projection. The equation did not require inputs of G_2 and N_2 but did require inputs of altitude, site index, and initial stocking.

$$d_2 = d_1 \exp[-(22.6132 + 0.00886 Altitude - 0.5208 SI - 0.0033 N_1)(\tfrac{1}{T_2^{1.0235}} - \tfrac{1}{T_1^{1.0235}})] \qquad (8.18)$$

$$\text{MSE} = 4.4166$$

Fitted MSE and graphical plots of residuals showed that equations based on relative basal area performed much better than equations based on sigmoid difference equations. With the approach based on relative basal area, it was the contribution of plot level G_2 and N_2 that led to the more precise projections for tree diameter.

Table 8.3: Fitting statistics for dbh equations based on sigmoid difference equations

No.	Equations	Number of parameters	MSE	Rank
8.5	Anamorphic Schumacher	2	5.0012	1
8.6	Polymorphic Schumacher	2	5.1824	8
8.7	Polymorphic Schumacher	1	5.3856	10
8.8	Anamorphic Chapman-Richard	2	5.0147	4
8.9	Anamorphic Weibull	2	5.0409	5
8.10	Anamorphic Gompertz	2	5.0734	7
8.11	Polymorphic Levakovic	2	5.3975	11
8.12	Anamorphic Levakovic	2	5.0050	3
8.13	Polymorphic Monomolecular	2	7.8429	12
8.14	Anamorphic Monomolecular	2	5.0013	2
8.15	Polymorphic Hossfeld	2	5.3008	9
8.16	Polymorphic Gompertz	2	7.9494	13
8.17	Polymorphic Gompertz	3	5.0568	6

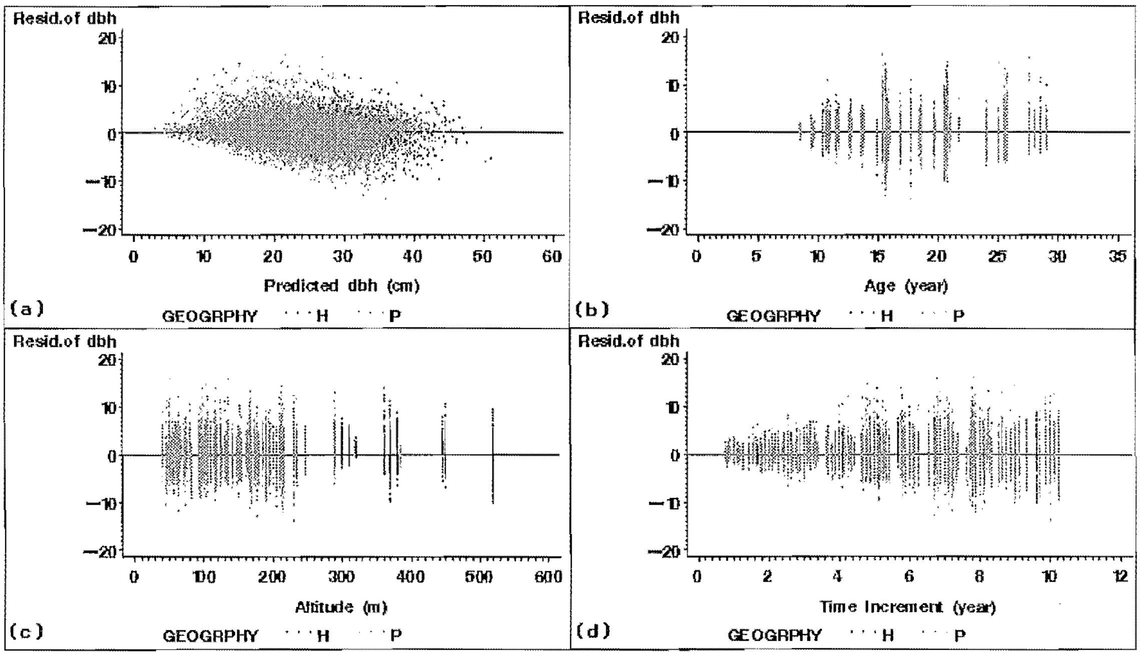


Figure 8.2: The fit of dbh projections for a model based on an expanded anamorphic Schumacher difference equation

To test the relationship between residuals of Equation 8.3 and explanatory variables, one pair of measurements for each tree and one plot in each stand was chosen randomly to obtain an autocorrelation-free data set. The test results with 3 864 trees (Table 8.4) showed that no explanatory variable was significantly correlated to the residuals. Only initial stocking was close to significance level at 5%. Its effect might be more important in older and denser stands and should be further investigated in future.

Table 8.4: Regression test results for residuals of dbh of individual trees

Dependent Variable	Independent Variable	DF	Parameter Estimate	Std. Error	T	Prob > |T|
Residual of dbh (Prob>F=0.25 R^2=0.015)	Intercept	1	-0.02434	0.3912	-0.062	0.9504
	Time increment	1	0.01879	0.0126	1.494	0.1353
	Altitude	1	-0.00001	0.0003	-0.019	0.9852
	Pruning height	1	-0.01800	0.0423	-0.426	0.6705
	GF rating	1	0.00627	0.0104	0.601	0.5480
	Initial stock	1	-0.00022	0.0001	-1.882	0.0599
	Site index	1	0.00388	0.0200	0.194	0.8464
	Error	3857				

8.4.2 Modelling of mortality of trees with Logistic regression

In the data set 3.5% of trees had died. It was hypothesised that tree death in the projection period was correlated to age, altitude, initial stocking, projection length, site index, GF rating, relative diameter or relative basal area, and the interaction between them. The logistic regression was fitted several times with different combinations of variables. The fitted results using full data showed that the best model for probability of tree death was Equation 8.21.

$$P(death) = (\exp(0.8104 - 2.8470\frac{d_i}{\bar{d}} + 0.00483\ Alt + 0.000191\ N_1 \times TimeInc - 0.1559\ SI)^{-1} + 1)^{-1} \tag{8.21}$$

where d_i = the initial diameter of i^{th} tree in a plot

$\bar{d}$ = arithmetic mean diameter in the plot

$\frac{d_i}{\bar{d}}$ = relative diameter (R_d) of i^{th} tree in the plot

$N_I \times TimeInc$ = the interaction between initial stocking and time increment

SI = site index.

To test the significance of those variables in Equation 8.21, logistic regression equations were fitted again with the same equation form as the above using a data set without autocorrelation and the results are listed in Table 8.5. The results indicated that relative diameter, altitude, the interaction between initial stock and time increment length, and site index were significantly correlated with tree death. Relative diameter was negatively correlated with tree death, which means the smaller trees in a plot were more likely to die. Trees died more often on poorer sites, higher altitudes, higher initial stockings with longer projection length (time increment). Relative basal area was tried but it was less effective than relative diameter in predicting tree mortality.

Table 8.5: Logistic regression results for tree mortality

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr>Chi-Square
Intercept	1	1.4931	1.1048	1.8266	0.1765
Relative diameter	1	-3.2188	0.4214	58.3382	0.0001
Altitude	1	0.0053	0.000844	39.4694	0.0001
$N_I \times$TimeIncrement	1	0.000179	0.000024	57.1324	0.0001
Site index	1	-0.171	0.0558	9.3771	0.0022
Error	3976				

8.5 PROJECTIONS AND DISAGGREGATIVE ADJUSTMENTS

Two projection models for stands and trees had been developed. In most cases, two projections from the two models might not produce exactly the same values for a plot or stand. Under the assumption that plot-level projections are more accurate than tree-level ones, disaggregative adjustments to projections of individual trees are needed so that an aggregation of adjusted values equals the plot-level projections.

To obtain projections for a list of trees in a plot, two steps had to be taken. The first step was to obtain projections using both stand and tree models and the second was to adjust the difference between the two projections.

For an initial list of n trees d_{11}, d_{12}, ... d_{1n} in a 0.04 ha plot, each tree may represent 25 trees on a one hectare basis ($n_{11} = n_{12} = \ldots = n_{1n} = 25$ and the total $N_1 = \sum n_{1i}$). The projected diameter and the represented tree number are written as d_{21}, d_{22}, ... d_{2n} and n_{21}, n_{22}, ... n_{2n}, respectively.

Step 1: Obtain two projections first using the stand model and then the tree model

With the stand level model established in the last chapter, future basal area, stocking per hectare and volume per hectare can be projected. The required equation forms are copied here:

$$G_2 = \exp[\ln(G_1)(T_1/T_2)^{\beta} + (\alpha_0 + \alpha_1 Alt + \alpha_2 (Alt - 250)X)/1000(1-(T_1/T_2)^{\beta})]$$

$$N_2 = (N_1^{c} + a(T_2^{b} - T_1^{b}))^{(1/c)}$$

$$V = \alpha G^{\beta} H^{\gamma}$$

With the tree level model, future surviving tree numbers can be projected with a tree mortality equation. Future diameter can be estimated with Equation 8.3. Height can be estimated with the height-diameter equation for regional scale and volume can be estimated with the tree volume equation described in Chapter 3. These equation forms are the as follows:

$$P(death) = (\exp(\alpha_0 - \alpha_1 Rd_1 + \alpha_2 Alt + \alpha_3 N_1 \times TimeInc - \alpha_4 SI)^{-1} + 1)^{-1}$$

$$n_{2i} = n_{1i}[1 - p(death)]$$

$$d_2 = \sqrt{\frac{1}{0.00007854}\frac{G_2}{N_2}[0.00007854 d_1^2 /(\frac{G_1}{N_1})]^{(T_2/T_1)^\beta}}$$

$$h = 1.40 + [0.695955 + 0.666983 T^{-0.5} - 0.106771\ln(SI) + (0.954201 + 0.000741 Alt)/d]^{-5}$$

$$v = \exp\{\alpha\ln(d) + \beta\ln[h^2/(h-1.4)] + \gamma\}$$

Step 2: Adjust projections of individual trees

adjust n_{2i} so that $\sum n'_{2i} = N_2$: $$n'_{2i} = n_{2i} + \frac{(N_2 - \Sigma n_{2i})n_{2i}d_{2i}^{-1}}{\Sigma n_{2i}d_{2i}^{-1}}$$

adjust d_{2i} so that $\Sigma g'_{2i}\, n'_{2i} = G_2$: $$d'_{2i} = d_{2i}\sqrt{\frac{G_2}{0.00007854\Sigma d_{2i}^2 n'_{2i}}}$$

adjust v_{2i} so that $\Sigma v'_{2i}\, n'_{2i} = V_2$: $$v'_{2i} = \frac{V_2}{\Sigma v_{2i}n'_{2i}} v_{2i}$$

The number of trees represented by each tree was adjusted with weightings of reciprocal of diameter. Larger adjustment was made for smaller trees, which is in accordance with Equation 8.21. Diameter adjustment is proportional to basal area and volume adjustment is proportional to volume.

8.6 Simulation Program for Model Implementation

There is little point in developing a growth and yield model unless it is to be used (Vanclay, 1994). Establishing a computer program provides users with an efficient means of applying the growth model.

8.6.1 The main features of the program

A computer simulation program, named CanSPBL, was designed to provide quantitative information of future production for both tree level and stand level. The simulation program integrates all model components described in both the previous chapter and this chapter and facilitates the use of the model. It was established as a flexible tool having capabilities in the following aspects:

a) It provides projections of tree dbh, total height, volume and surviving probability for all individual trees in a plot. The detailed information based on individual trees is especially useful for harvest planning.

b) It simulates one plot or one stand growth in a number of projection times as the user desires, which is similar to model CANTY, one of the most used forms in New Zealand. The components of MTH, basal area per hectare, stems per hectare, volume per hectare, and diameter distributions are provided.

c) It produces future projections of multiple stands or plots simultaneously given the recent measurements available. Broad information for all targeted stands in a forest or forests is needed to provide yield estimates as inputs for estate modelling and resource evaluation. STANDPAK (Whiteside, 1990) is a versatile system for a single stand but it may be inefficient in projecting future yields for a company owning thousands of stands, all varied in site, crop type, age and other conditions.

8.6.2 The construction of the program

Excel 97 was selected as a host for the projection program for its sheet utility in operating data, its availability among most users and the preference of specified users in SPBL. Microsoft Office 97 provides Visual Basic as a programming tool to define and execute a series of actions. Inputting initial values by copying data from an existing worksheet or a database is preferable to letter by letter keyboard operations when multiple units (stands, or plots or trees) are considered at a time.

There are one hidden sheet, two forms and three standard modules in the program file of CanSPBL.xls. One form is for the introduction and the other for displaying a help text document. The procedures in the program modules include opening up a new user file, defining a toolbar containing five command buttons, formatting sheets and headers, reading user inputs, checking input validity, computing and outputting projections into specified cells, and handling errors. A user file (workbook) is created with and separated from the supporting file (CanSPBL.xls); a user saves only user data when the file is saved. The user interfaces with two main working sheets in a newly opened user file. One is for stand or plot level, and the other is for trees in a plot.

The program is simple, flexible, and portable. Its simplicity is reflected by its visual interface with users and the ease of entering data in Excel sheets. Users can specify the projection length and projection times to reach the desired simulations. The size of the program is less than 400 kilobytes.

8.6.3 Using the program

Two files are packed in this program and they are stored on the appended floppy diskette. The file CanSPBL.xls is the main program and it can only be run in Excel 97. The file Readme.txt is a guide for installation.

A document is provided within the program and it can be accessed by clicking the HELP command button. The various aspects of using the program have included the user interface, the procedures to obtain projections for stands or plots, the procedures to obtain projections for trees in a plot, dbh distribution, the range of initial values, and the accuracy and limitations of the program.

8.7 DISCUSSION

The reason that relative basal area based equations fitted tree diameter so well may be that an equation employing such variables as T_1, T_2, G_1, G_2, N_1, N_2 restricted the tree-level variations of error in projecting diameters. Now that the projection equations of basal area and stocking have been established, they could be used to contribute to obtaining reliable diameter estimates at a tree level. This method can be applied to other species in other region if a stand level model can be established reliably.

It should be pointed out that the disaggregative adjustments applied in this study are conceptually different from the disaggregation in other studies. The disaggregation in Pienaar and Harrison (1988), Ritchie and Hann (1997) and Knowe *et al.* (1997) is to allocate the increment built up during a period of time (T_2-T_1). The disaggregative adjustments applied in this study are to allocate the projection differences between stand and tree models at the same projection age (T_2). Therefore the conclusions draw by Ritchie and Hann (1997) and Knowe *et al.* (1997) may or may not apply to this study. The employment of a logistic equation for mortality in this study reduced the subjectivity involved when determining mortality of a given tree during a projection interval proportionally to its inverse relative basal area (Pienaar and Harrison, 1988).

There may be three types of relations between a stand model and a tree model: (i) mutually independent, (ii) completely compatible and (iii) one dependent on the other. For the first type, each model is established separately and no adjustment will be made when predictions from the two models are different. For the second type, both a stand model and a tree model may be established simultaneously so that the predictions from the two models are the same (Zhang *et al.*, 1997). For the third type, if the tree model is considered to be true then stand statistics should be obtained with the integration of individual trees but if the stand model is considered more accurate then the predictions for individual trees should be adjusted to allow the same projections from the two models. Stand models are generally believed accurate in New Zealand (Garcia, 1991; Goulding, 1994) and in other countries (Somers and Nepal, 1994). This study adopted the same view and disaggregative adjustments were made to individual-tree projections, but the assumption should be justified in future studies. The second type above seems more desirable but further research is needed to improve tree model for longer projections (Zhang *et al.*, 1997).

A simulation program that is based on operating systems (Windows or Macintosh) would be preferable for users without Excel 97. The facility of transferring data from or integrating with a PSP database would increase its usefulness in selection of targeted records for projections. This will be studied in future.

8.8 CONCLUSIONS

The two necessary components of diameter and survival for completion of an individual-tree modelling system were built. For the projection of diameter of individual-trees, an approach based on relative basal area led to a best fit in comparison with many sigmoid difference equations. Basal area and stocking at a plot level contributed to the reliability of diameter projections at tree levels. For projections up to 10.5 years, 90% of residuals from a relative basal area based equation were within ±3.1 cm.

A logistic regression procedure showed that relative diameter, altitude, the interaction between initial stocking and time increment length, and site index were significantly correlated to tree mortality. Relative diameter was negatively correlated with tree death, which means that smaller trees in a plot were more likely to die. Trees died more often on poorer sites, higher altitudes, and higher initial stockings with longer projection lengths.

A computer simulation program was developed for model implementation at both stand and tree level. It produces estimates of multiple stands or plots simultaneously given recent measurements. It is simple and flexible but only applicable in Microsoft Excel 97.

CHAPTER 9

JUVENILE GROWTH MODELLING

9.1 INTRODUCTION

Juvenile growth modelling for plantations of *Pinus radiata* concentrates on the state from immediately after planting to around age five years before thinning. Juvenile growth is sensitive to environmental and silvicultural operations and managers are interested in quantitative estimates of these effects. The objectives of juvenile growth modelling are to identify the main variables and predict growth responses to different sites and different silvicultural options.

Juvenile growth can be simply expressed as a function of conditions of genetics, seedling status, sites, and treatments at any stage. Seedling quality can be described physically and morphologically and it can be changed by genetic improvement, nursery regimes and stock handling (Genetics and Tree Improvement Research Field, 1987; Menzies, 1988). Site productivity is determined by location, soil, weather and other factors (Hunter and Gibson, 1984). The micro-environment can be changed by such field management options as land clearing, cultivation, fertilisation, and weed management (Balneaves, 1982; Hunter and Graham, 1982; Mason *et al.*, 1995; Mason *et al.*, 1996).

In order to reflect growth responses fully for the different options yield-age equations were employed by modellers (Belli, 1987; Belli and Ek, 1988; Mason, 1992; Zhao and Mason, 1996; Mason and Whyte, 1997). Difference forms can hide effects behind combinations of initial time T_1 and initial yield Y_1. Belli (1987) and Belli and Ek (1988) modelled growth and survival for conifer plantations in the Great Lakes Region of the United State of America. Mason (1992) created a juvenile growth model for radiata pine plantations in the central North Island of New Zealand and analyses of variance for equation coefficients created with non-linear regression for each individual plot were used to identify significant variables. The study described here is a continuation of the study by Zhao and Mason (1996, see Appendices) of juvenile growth in Canterbury, New Zealand. Zhao and Mason (1996) transformed equations into linear forms and the parameter prediction equations were tested and established simultaneously. Annual rainfall, weeding and nitrogen fertiliser were identified and introduced into the model of mean height and ground level diameter. The three procedures of principal component analysis to select environmental variables, linear multivariate regression to test parameter hypotheses with an autocorrelation-free data set and non-linear regression to recalibrate parameter estimates with the full data set were very effective. Both linear and non-linear procedures were conducted by using the direct response variable so the two analyses were more compatible. One temporary plot measurement could be used with linear regression in the system, but at least three measurements in a plot were needed with the non-linear equation to obtain the parameter estimates for the plot. It was claimed that more data covering a wider range sites were needed, however, to re-calibrate parameters and to predict growth more precisely in the wide range of site conditions represented in the model.

To refine the juvenile growth model in this study, it was determined that the model components were to be modified and some new temporary plot measurements were to be added to the experimental data set. With the support of SPBL in November 1997, thirty-one temporary plots were sampled within the company's estate. The main components of the juvenile growth model were to be mean top height, sectional area at ground level, basal area, and survival. These components reflect the features of juvenile growth and are compatible with models of older growth.

9.2 DATA DESCRIPTION

Two sources of data were available for the study: experimental records and passive temporary plot measurements. Experiments were set mainly in the forests of Carter Holt Harvey and temporary plots were located in the forests of SPBL.

In order to relate growth equations with potential environmental variables accounting for growth differences between experiments, altitude and average annual rainfall during the experimental periods were collected. They were the easily available and potentially important explanatory variables (Zhao and Mason, 1996). The correlation coefficient between altitude and annual rainfall was 0.5.

9.2.1 Experiments

There were eleven sets of experimental data available which were serial uniformity trials used to examine the response of radiata pine to cultivation, weeding and fertilisation during establishment. The experiments were located in four main regions: Balmoral Forest, Eyrewell Forest, Hanmer Forest and Mcleans Island Forest. Nine of the experiments were established in the 1970's by the New Zealand Forest Research Institute (FRI) and two others were set up in 1983 by Euan Mason and Patrick Milne. Height and diameter at ground level (dgl) measurements were available mainly for up to four years. Table 9.1 lists the information for each experiment.

Weed Control was designed as a binary variable but its values were adjusted to a range between 1 and 0 in the modelling process, based on the discovery from some experiment documents that actual weeding was poorly carried out. A value of one was given to the weeded treatments in experiments C573/1 and C573/2, which were established by Mason and Milne in 1983. Fertiliser nitrogen (N), phosphorus (P), potassium (K) and boron (B) were quantitative variables, with volume varying according to amounts added to each tree. Cultivation was not significant in preliminary analyses so the levels of cultivation were not distinguished.

Table 9.1: Summary information for each individual experiment

Experiment	Location	Age at which data available	Cultivation	Fertilisation Layout	Weed control	Initial stocking	Altitude	Annual rainfall during trial
[1]C393	Hanmer	0,2,4	0, 1	N×P (3×4)	0	1670	400	1446
[1]C394	Balmoral	0-4	0	0, N+P+B	1	1667	200	666
[1]C396	Eyrewell	0,2,4	0	N×P (3×4)	0	1250	140	933
[1]C404	Balmoral	0,2,4	0,1	N×P (3×4)	0	1667	290	647
[1]C451	M.Lean.Isl	2,4,6	0	0, P+B	1	1764	75	760
[1]C507/1	Balmoral	1-6	0	0, N+P	0,1	625	295	655
[1]C507/2	Eyrewell	0-3	0	0, N+P	0,1	667	150	887
[1]C515	Balmoral	0-4	0	0, N+P+	1	1000	260	629
[1]C516	Balmoral	0-4	0	N×P×K×S (2^4)	1	1000	250	629
[2]C573/1	Eyrewell	0-5	0,1	0, N+P+K+B	0,1	1250	135	786
[2]C573/2	Balmoral	0-4	0,1	0, N+P+K+B	0,1	1250	240	681

[1] Established by Forestry Research Institute (FRI) in 1970's, [2] established by Euan Mason *et al.* in 1983

9.2.2 Temporary plot measurements

During the period of this study, thirty-one temporary plots were sampled and measured in November 1997 within the boundary between the Waimakariri River and the Rakaia River in SPBL's estate. Thirteen plots were located on plains, seven on foothills and eleven on coastal sands. Age was distributed evenly on each geographical type ranging from 1.45 to 7.45 years old, which represented an establishing period from 1990 to 1996. None of the plots had been thinned. Seedlings used by SPBL during the period were rating as GF 16 stock. Weeding was carried out mostly during the first two years and actual weed coverage was recorded on site. Plot area was 0.04 ha. Ground level diameter and height were all measured and dbh was measured for all trees exceeding 1.40 m in height. The number of planted trees and dead trees were counted and recorded. Initial height and initial basal area at age zero were not available for all temporary plots.

Table 9.2: Summary statistics for temporary plots

Variable	Frequency	Mean	Std Dev	Min.	Max.
Weed Coverage	31	0.5	0.2	0.1	0.9
Altitude (m)	31	175	197	4	640
Rainfall (mm)	31	754	148	600	1000
T (year)	31	4.51	2.16	1.45	7.45
Mean top height (H) (m)	31	4.01	2.58	0.43	9.19
Basal area/ha at ground level (G_{GL})(m^2)	31	8.732	10.317	0.034	44.473
Basal area/ha at breast height (G) (m^2)	31	4.224	5.874	0.000	23.609
Survival (S)	31	0.87	0.14	0.47	1.00
Live stocking (N)	31	1051	271	475	1500

The 31 temporary plots were to be used to bridge basal area with sectional area at ground level. In all experiments, measurements were only available for either diameter at ground level or diameter at breast height but not for both. Diameter at ground level was replaced by diameter at breast height when trees reached the age 5 or older. For modelling sectional area at ground level, mean top height and survival, temporary plots younger than or equal to five years of age were used to match with the age of the experimental data since the very few points with higher ages could cause extremely large kurtosis of model errors.

The number of the sampled temporary plots was much smaller than the number of plots derived from experimental data. The temporary plots represented trees in SPBL's estate.

9.3 Equations and Use

Yield-age functional forms were adopted in this study. Sigmoid curves cannot be expected during the juvenile growth period. The following types of equations were employed:

$$H_T = H_0 + \alpha T^{\beta} \tag{9.1}$$

$$G_{GL.T} = G_{GL.0} + \alpha\ N_0 T^{\beta} \tag{9.2}$$

$$G_T = \alpha G_{GL.T}[1 - \exp(-\beta G^{\gamma}{}_{GL.T})] \tag{9.3}$$

$$S_T = e^{\alpha T^{\beta}} \tag{9.4}$$

where H_T = mean top height at age T

H_0 = initial MTH at age zero

$G_{GL.T}$ = sectional area at ground level at age T

$G_{GL.0}$ = sectional area at ground level at age zero

N_0 = initial stocking just after planting

S_T = survival at age T

G_T = basal area at age T

α, β, γ = parameters in the equations

Equation 9.1 had been a good basic form for height or diameter, Equation 9.2 for sectional area at ground level and Equation 9.4 for survival modelling (Belli, 1987; Belli and Ek, 1988; Mason, 1992; Zhao and Mason, 1996; Mason, 1997). When the parameters in Equation 9.1 are assumed a linear combination of explanatory variables and when the first parameter α in the equation is written as e^{a} then the equation becomes Equation 9.5 that can be transformed with natural logarithms into a linear equation (Equation 9.6).

$$H_T = H_0 + e^{(a_0+a_1V_1+a_2V_2+\ldots a_nV_n)} T^{(b_0+b_1V_1+b_2V_2+\ldots+b_nV_n)} \tag{9.5}$$

$$\ln(H_T - H_0) = a_0 + a_1V_1 + a_2V_2 + \ldots a_nV_n + b_0\ln(T) + b_1V_1\ln(T) + b_2V_2\ln(T) + \ldots + b_n V_n \ln(T) \tag{9.6}$$

where n = the number of total explanatory variables

V_1, V_2... V_n = explanatory variables

a_1, a_2...a_n, b_1, b_2,...b_n = parameters

Similarly, Equation 9.2 can be transformed into a linear model (Equation 9.7). Tests of hypotheses can then be done by using linear regression with direct response as a dependent variable rather than using analyses of variance of the parameter estimates.

$$\ln\left(\frac{G_{GLT}-G_{GL0}}{N_0}\right)=a_0+a_1V_1+a_2V_2+\dots a_nV_n+b_0\ln(T)+b_1V_1\ln(T)+b_2V_2\ln(T)+\dots+b_n\,V_n\ln(T) \quad (9.7)$$

For the relationship between basal area and sectional area at ground level, a linear form could fit for the most part except at the very stage at which trees started to reach breast height. Equation 9.3 was proposed in the study to represent the relationship between basal area and sectional area at ground level. It is a form of modified Weibull equation and has no asymptote. Basal area is zero when sectional area at ground level (G_{GL}) is zero, and is proportional to G_{GL} when G_{GL} approaches a large value. Parameter α is the proportion, the ratio of basal area to ground level sectional area.

9.4 Methods

Analyses of variance or analyses of covariance for individual experiments were conducted according to the designs of each experiment. In an analysis of covariance, initial value immediately after planting (initial height or initial diameter at ground level) was taken as covariate and treatments were taken as classification variables. Analysis of variance was used for two experiments in which no initial measurements were available. Using procedure GLM in SAS (SAS Institute Inc., 1992), one analysis was done for each year's measurements of height and diameter at ground level or diameter at breast height.

The growth model was developed using two stages of regression. One was linear multivariate regression to test hypotheses, and the other was non-linear regression to recalibrate parameter estimates with the full data set.

To test the parameters in the transformed linear model simultaneously, stepwise regressions were employed using direct response variables. Linear regression could deal with a number of explanatory variables in an effective and unique solution while non-linear regression could lead to mathematical difficulties if too many variables were involved. Autocorrelation-free data were used in the procedure to eliminate dependence of errors due to repeated measurements.

To recalibrate the parameter estimates in the selected model, non-linear procedures were run using the full data set. This procedure was aimed to fully use all available information and to enable the model to be more representative. Non-linear regressions were run by including the tested significant variables from the above linear regression. It was easy to get non-linear procedure (SAS Institute Inc., 1990) to converge after a number of variables were filtered by the linear regression procedure.

Four graphical plots of residuals were displayed for each of the model components and they were prediction, age, altitude, and rainfall.

9.5 RESULTS

9.5.1 Analyses of individual experiments

Analyses of variance or analyses of covariance were conducted for each experiment individually and a summary of the results of hypothesis testing is shown in Table 9.3.

There were four trials designed to test cultivation and only one of them revealed a significant effect on growth of height and two were significant on diameter. Weeding, however, was positively significant in all four designed experiments for both height and diameter. Survival often did not change significantly with the treatments but two trials showed survival reduction with weeding due to Velpar poisoning and two others with fertilisation.

Table 9.3: Summary of hypothesis testing results for individual experiments

Experi-ment	Age Tested	Cultivation h	Cultivation d	Cultivation s	Fertilisation h	Fertilisation d	Fertilisation s	Weeding h	Weeding d	Weeding s
C393	4	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.			
C394	4				NPB+	NPB+	n.s.			
C396	4				P -	N×P -	n.s.			
C404	4	+	+	n.s.	N +	N+	n.s.			
C451	6				B -	n.s.	n.s.			
C507/1	6				NP +	NP +	NP-	+	+	n.s.
C507/2	3				NP +	NP +	NP-	+	+	-
C515	4				NP +	NP +	n.s.			
C516	4				N+ K+ N×S-	N+ K+ N×S-	n.s.			
C573/1	5	n.s.	+	n.s.	n.s.	n.s.	n.s.	+	+	n.s.
C573/2	4	n.s.	n.s.	n.s.	NPKB+	NPKB+	n.s.	+	+	-

Notes:

h, d, s = height, diameter (dbh or gld), and survival, respectively

blank = not designed to test

\+ = positively significant response

\- = negatively significant response

n.s. = not significant even at a level of 5%

NP+ = positively significant N & P combination

N×S- negatively significant interaction between N and S

The fertilisation involved was mainly nitrogen, phosphorous, potassium, and boron. Nitrogen or combinations of nitrogen with other elements significantly increased height and diameter growth in seven out of ten experiments in which nitrogen was a factor. Nitrogen on its own was positively significant in two of four factorial designed experiments and not significant in the other two. Phosphorous was involved in all eleven experiments and it was negatively significant in one of the four designed trials and not significant in the other three. Potassium was involved in four experiments and its significance was shown only in experiment C516. Boron was involved in four experiments and it showed a positive significance and a negative significance in two of the designed trials.

Covariate of diameter at ground level immediately after planting was significant for height and diameter at age four in trials C393 and C404.

Weeding and nitrogen fertilisation most commonly increased height and diameter growth. The exact magnitudes of the effects on a regional scale rather than in individual experiments could not be identified without further growth modelling.

Such treatments in the experiments as legume treatments in C507/1, C507/2, re-fertilisation in C451, C515 and C516 were uncommon so they were dropped from modelling process. Nine sets of experimental data were left for the following growth modelling.

9.5.2 Modelling mean top height (H)

The results of testing for explanatory variables in the height model (Equation 9.6) with stepwise linear regressions were that rainfall and weeding entered the model significantly after age entered. Table 9.4 shows the regression results from an autocorrelation-free data set. Keeping the above variables and setting the parameter estimates as starting values, non-linear procedures with the full data set of 164 observations produced the new parameter estimates that are listed in Table 9.5. The resulting equation form was Equation 9.8. Rainfall and weeding were important factors increasing growth. The mean square error (MSE) for Equation 9.8 was 0.053 that was 39% of MSE before the two variables were introduced into the model. The residual patterns against predicted value, age, altitude and average annual rainfall are shown in Figure 9.1. There was little apparent bias. Residuals were all within ±0.8 m.

$$H_T = H_0 + e^{(a_0+a_1 RAINFALL)} T^{(b_0+b_1 WEEDING)} \quad (9.8)$$

where T = age

H_T = mean top height at age T

H_0 = initial MTH at age zero and it was replaced by 0.3 m (the mean value for all available H_0) when H_0 was unknown

a_0, a_1, b_0, b_1, = coefficients shown in Table 9.4

Table 9.4: Linear regression results for H obtained with autocorrelation-free data

Variable	DF	Parameter Estimate	Std. Error	T	Prob>|T|
Intercept	1	-2.071478	0.13091	-15.823	0.0001
Ln(age)	1	1.585267	0.08614	18.404	0.0001
Rainfall	1	0.000647	0.00015	4.434	0.0001
Weeding × ln(age)	1	0.435471	0.10977	3.967	0.0002
Error	67				

$R^2 = 0.896$

Table 9.5: Non-linear fitting results for H obtained with a full data set

Variable	Parameter	Parameter Estimate	Asymptotic Std. Error	95 % Confidence Interval Lower	Upper
Intercept	a_0	-1.896823	0.074570	-2.044092	-1.749553
Ln(age)	b_0	0.0006263	0.000044	0.000540	0.000713
Rainfall	a_1	0.3721068	0.030835	0.311211	0.433003
Weeding	b_1	1.4783679	0.049660	1.380293	1.576442

Total observations = 164

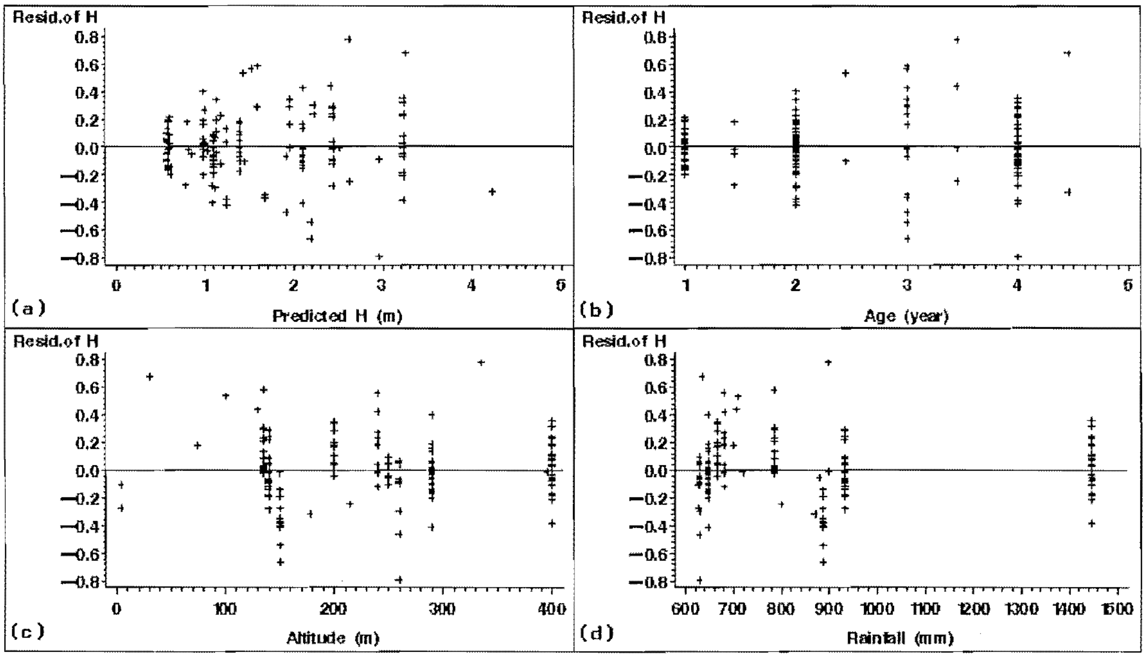


Figure 9.1: The residual pattern of MTH of juvenile growth

9.5.3 Modelling ground level sectional area per hectare (G_{GL})

Sectional area at ground level was fitted with the same procedures as mean top height. Significant variables entered into the G_{GL} model (Equation 9.7) in stepwise linear regression procedures were also age, annual rainfall, and weeding. Table 9.6 lists the linear regression results, which were obtained using the autocorrelation-free data set. Table 9.7 shows the re-calibrated parameter estimates for Equation 9.9 obtained with non-linear procedures using the full data set. The MSE was 0.167, 15% of MSE before rainfall and weeding introduced into the model. Figure 9.2 shows the residual pattern. No apparent bias appeared. Residuals were all within ± 1.5 m^2.

$$G_{GL.T} = G_{GL.0} + e^{a_0} N_0 \, T^{(b_0 + b_1 RAINFALL + b_2 WEEDING)} \qquad (9.9)$$

where T is age, $G_{GL.T}$ is sectional area at ground level at age T, $G_{GL.0}$ is sectional area at age zero (mean value 0.3 was given if $G_{GL.0}$ was unknown), N_0 is initial stocking just after planting, a_0, b_0, b_1, b_2 are coefficients.

Table 9.6: Linear regression result for G_{GL} obtained with autocorrelation-free data

Variable	DF	Parameter Estimate	Std. Error	T	Prob > |T|
Intercept	1	-9.618354	0.09311	-103.303	0.0001
ln(age)	1	1.673664	0.20845	8.029	0.0001
Weeding × ln(age)	1	0.613436	0.13966	4.392	0.0001
Rainfall × ln(age)	1	0.000962	0.00017	5.548	0.0001
Error	67				

R^2=0.9

Table 9.7: Non-linear fitting results for G_{GL} obtained with a full data set

Variable	Parameter	Parameter Estimate	Asymptotic Std. Error	95 % Confidence Interval	
				Lower	Upper
Intercept	a_0	-9.651174	0.124657	-9.897361	-9.404987
Ln(age)	b_0	1.9071179	0.101847	1.705979	2.108257
Rainfall	b_1	0.0007883	0.000032	0.000726	0.000851
Weeding	b_2	0.6587929	0.031397	0.596787	0.720799

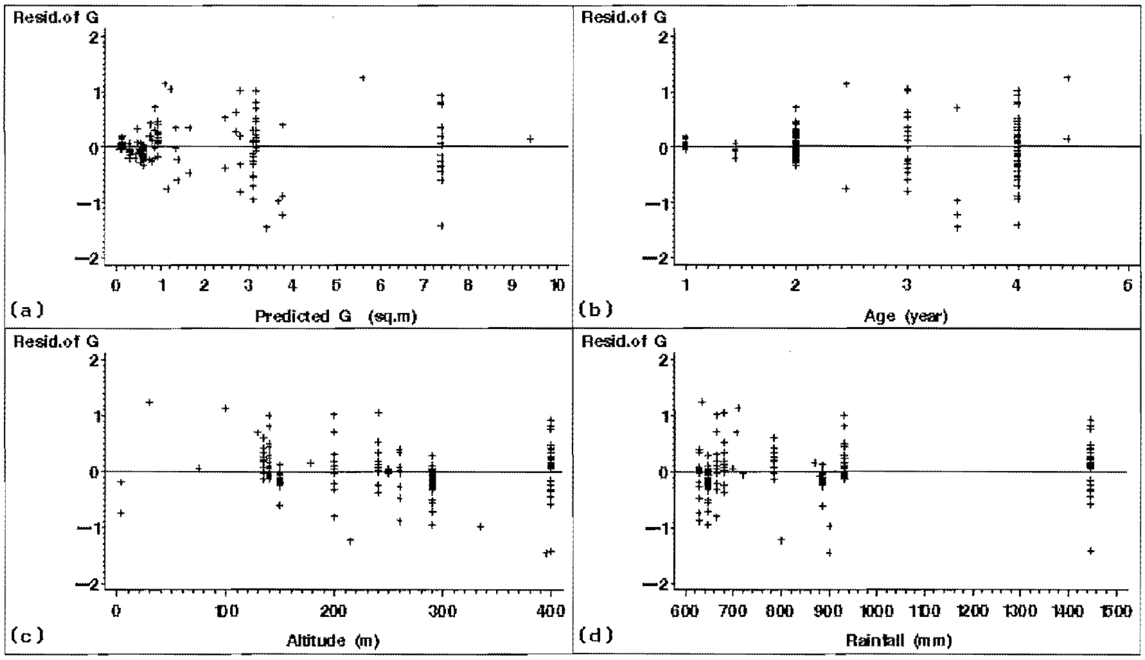


Figure 9.2: The residual pattern of ground level sectional area of juvenile growth

9.5.4 Modelling survival (S)

Equation 9.4 was used for modelling survival. Its expanded model can be written as Equation 9.10.

$$S_T = e^{(a_0+a_1V_1+a_2V_2+...+a_nV_n)T^{(b_0+b_1V_1+b_2V_2+...+b_nV_n)}} \tag{9.10}$$

where n was the number of total explanatory variables, V_1, V_2... V_n were explanatory variables, a_1, a_2... a_n, b_1, b_2... b_n were parameters in the model.

Parameter estimates were created for each plot with non-linear procedures first. The linear regression of the parameter estimates against independent variables showed that the second parameter β was correlated to nothing significantly. When parameter β was estimated with a full data set as a constant b_0, the model then could be transformed with logarithms into linear model (Equation 9.11). After parameters were tested for the model with linear regression of survival against explanatory variables directly, annual rainfall was indicated a significant variable (Table 9.8). Results of final re-calibration of parameter estimates for Equation 9.12 with non-linear procedures using full data set are listed in Table 9.9. The effect of rainfall on survival was not as significant as that on MTH and basal area. Only a 7.6% reduction in MSE was gained after the introduction of rainfall into the model. Weeding did not improve survival significantly, which had been shown in the analyses for individual experiments. Figure 9.3 shows the model's residual pattern.

$$\ln(S_T) = a_0T^{b_0} + a_1V_1T^{b_0} + a_2V_2T^{b_0} + ... + a_nV_nT^{b_0} \tag{9.11}$$

$$S_T = e^{(a_0+a_1RAINFALL)\,T^{b_0}} \tag{9.12}$$

where n was the number of total explanatory variables, V_1, V_2... V_n were explanatory variables, a_1, a_2, b_1, b_2,...b_n were parameters in the model.

Table 9.8: Linear regression results for S obtained with autocorrelation-free data

Variable	DF	Parameter Estimate	Std. Error	T	Prob>|T|
T^{05}	1	-0.05869700	0.012296	-4.774	0.0001
Rainfall $\times T^{0.5}$	1	0.00003522	0.000013	2.787	0.0069
Error	67				

R^2=0.45

Table 9.9: Non-linear fitting results for survival obtained with a full data set

Variable	Parameter	Parameter Estimate	Asymptotic Std. Error	95 % Confidence Interval Lower	Upper
Intercept	a_0	-0.047931	0.009888	-0.067457	-0.028405
Rainfall	a_1	0.0000265	0.000008	0.000011	0.000042
Intercept	b_0	0.5105184	0.177992	0.159016	0.862021

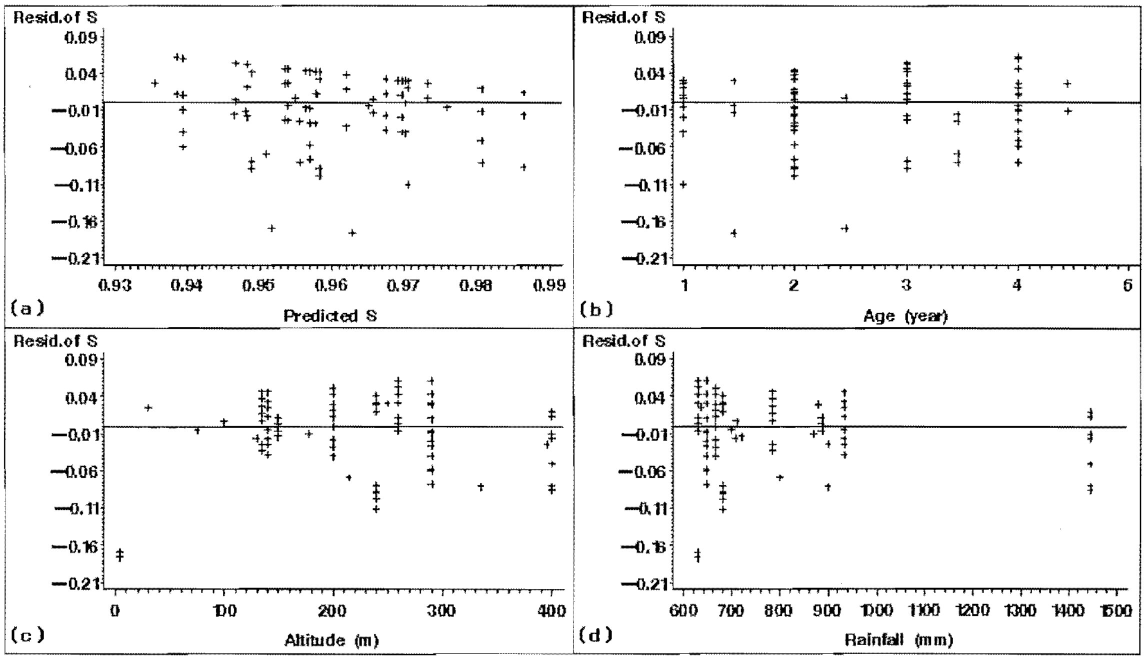


Figure 9.3: The residual pattern of survival of juvenile growth

9.5.5 Modelling basal area per hectare (G)

Only 31 temporary plots were available for this model. Ground level diameter (gld) was mostly measured for experimental data but gld was not available if diameter at breast height (dbh) was.

Table 9.10 shows parameter estimates for Equation 9.3. The MSE was 0.455 when the fitting procedure converged. Residuals were all within ± 1.5 m^2. Figure 9.4 shows the residual pattern of the model. Little bias was apparent but the effects of rainfall and altitude should be investigated with more data.

Table 9.10: Non-linear fitting results for G obtained with temporary plots

Parameter	Parameter Estimate	Asymptotic Std. Error	95 % Confidence Interval Lower	Upper
α	0.556839	0.01297	0.53028	0.58340
β	-0.14179	0.09145	-0.32912	0.04555
γ	1.060254	0.28375	0.47902	1.64149

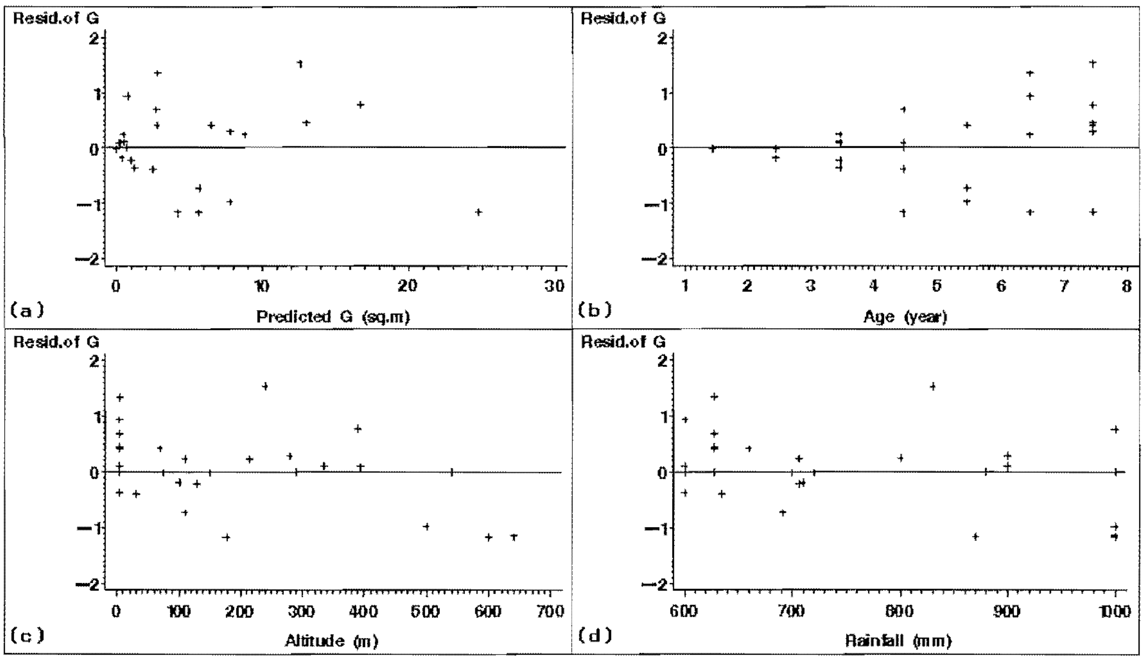


Figure 9.4: The residual pattern of basal area of juvenile growth

Figure 9.5 shows the trend of basal area (at breast height) versus sectional area at ground level. A linear form could fit most of the data except at very young stages at which trees started to reach breast height. The MSE with a linear regression was 0.932, which is twice the MSE with Equation 9.3.

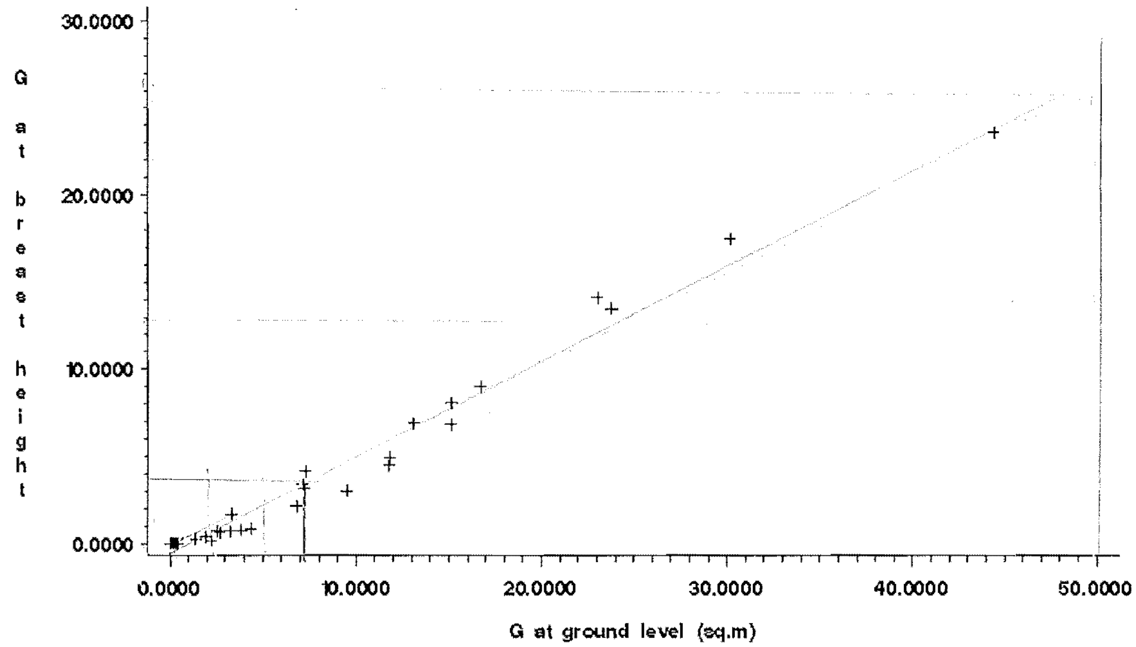


Figure 9.5: The trend of basal area versus sectional area at ground level

9.6 Discussion

The regression procedures of linear regression to test parameters with an autocorrelation-free data set, and non-linear regression to recalibrate parameter estimates with the full data set, were very effective for juvenile growth modelling. An equation that could not be transformed into a linear form, however, might not be modelled easily with these methods.

The model was more improved with the introduction of average rainfall during the trial period than with the introduction of altitude. Altitude does not represent fluctuation of weather variation in different years. Both rainfall during the trial and weeding could describe growth differences within and among experiments. Nitrogen fertiliser was not significant in this study but it was significant when only seven experiments were used (Zhao and Mason, 1996). Its effect was too weak to enter into the new model for which more data were included in this study. GF rating might significantly affect growth but it was not detectable in this study. It should be investigated further in future.

This juvenile growth model built with the addition of newly taken temporary plot measurements was expected to suit wider conditions than the model of Zhao and Mason (1996). But the heavy weight from experimental data set means prediction differences from the two models might not be great. Considering the model was for the whole Canterbury region, more data were needed to improve the model's predictions. The model was not validated due to a lack of sufficient data.

9.7 CONCLUSIONS

This study aimed to model the growth of radiata pine aged younger than 5 years in Canterbury. Nine sets of experimental data and thirty-one temporary plots were used to build a juvenile growth model with the components of mean top height, sectional area at ground level, basal area, and survival. Regression procedures revealed that mean top height and basal area were significantly related to average rainfall during trial periods and weeding practice. Larger trees were produced from higher rainfall and with weeds eliminated. Higher rainfall could reduce mortality significantly but affected survival less than mean top height and basal area. Killing weeds did not increase survival and this was consistent with findings from analyses of individual experiments.

CHAPTER 10

GENERAL DISCUSSION

A broad research area was involved in the study, beginning with database design and ending with the construction of a simulation program for end-users. Some particular discussion provided in each chapter for each separate study will not be repeated here. New features in the study, limitations of the established models, some relevant specific points and future research in modelling are discussed.

10.1 NEW FEATURES IN THE RESEARCH

10.1.1 Improvements in model CanSPBL in comparison with model CANTY

An existing model CANTY is a stand model with 3 state variables consisting of mean top height, basal area per hectare, stems per hectare, and output of volume per hectare. Model CanSPBL established in this study aims at multi-resolution applications and it includes the above four components, as well as diameter distribution, tree diameter, tree mortality and tree height (height-diameter equation).

Residuals of mean top height and basal area with CANTY showed a close correlation with altitude. The effect of altitude was formulated in this study to differentiate plains and hills. Incorporation of altitude in the selected equation reduced mean square error by 17% for MTH and 41% for basal area/ha.

That large numbers of plots and trees were used to build model CanSPBL was one of the features of this study. This made the model more representative and reliable.

10.1.2 Modelling of individual trees

To model diameter of individual-trees, an approach derived from relative basal area projection equation resulted in the best fit. This approach is a further development based on the work by Clutter and Jones (1980) and Pienaar and Harrison (1988). Borders and Patterson (1990) compared three methods to project stand tables and it was found that the relative basal area method of Pienaar and Harrison (1988) was superior to a Weibull diameter distribution method and a percentile-based method. Allocation of mortality for each diameter-class was quite subjective with the method of Pienaar and Harrison (1988), however. Mortality of a given tree during a projection interval was simply assumed proportional to its inverse relative basal area. This study here employed two steps to allocate mortality for a given tree. The first step was to project tree mortality using a logistic mortality equation which was sensitive to relative diameter, site index, altitude, initial stocking and projection length. The second step was to adjust the tree mortality so that projections from both a stand model and a tree model were consistent. Projection differences between stand and tree models at age T_2 were usually much smaller than the change of stocking per hectare from initial age T_1 to projection age T_2, so the potential bias with the subjective adjustment at this second step was far less than with the method of Pienaar and Harrison (1988). In this study the relative basal area projection equation was directly transformed to a diameter projection equation. A newly proposed formulation of diameter projection equation in this study fitted equally well to the equation of Pienaar and Harrison (1988).

An approach derived from relative basal area projection equation has shown the best fit in contrast with many sigmoid difference equations. Basal area and stocking projected at stand level model contributed to the reliability of diameter projections of individual trees. This method can be applied to other species and in other regions if a stand level model can be established reliably.

10.1.3 Development of a stand model

Stand models are generally developed with statistics of plots. To obtain stand statistics with random plots, estimates are not biased for such variables as mean height, basal area/ha, stems/ha, and volume/ha but biased for variance, percentiles, and extreme values such as maximum and minimum diameter due to the spatial correlation (Garcia, 1991). In this study the model components of MTH, basal area, stems per hectare, and volume per hectare were developed with plot statistics and showed no bias for the projections at a stand level. For modelling diameter distributions, diameter deviation and maximum diameter were estimated with stand level statistics. This approach can rarely be seen in other diameter distribution models for stands.

The bias of estimating stand variance with mean of plot variance has long been discussed in Garcia (1991) and Whyte and Woollons (1992), and this study provided an opportunity to display the magnitude of the bias. This study showed with 402 stand measurements that the bias of underestimation was not serious for most stands but was for a few stands.

10.1.4 A simulation program to project multi-stand simultaneously

A computer simulation program was developed in this study to produce future projections of multiple stands or plots simultaneously given the recent measurements. The program allows users to project their production over a whole region and obtain information for estate modelling more efficiently than any other programs used in New Zealand. STANDPAK (Whiteside, 1990), the most ubiquitous stand modelling package in New Zealand, projects only stand-level statistics for a single stand. So does CANTY.

10.1.5 Improvement in Juvenile growth model

To refine the juvenile growth model of Zhao and Mason (1996), this study replaced mean height and diameter with mean top height, sectional area at ground level and basal area that are more comparable with models for older growth. The newly proposed equation to estimate basal area with sectional area at ground level fitted the study data very well.

For modelling juvenile growth and yield with yield-age equations, this study demonstrated the use of temporary plot measurements with experimental data. At least three measurements in a plot were needed to test explanatory variables with the approach used by Mason (1992), and Mason and Whyte (1997), but plots with only one measurement could be well used for modelling height and basal area using the procedures described in Zhao and Mason (1996) and in this study.

10.2 LIMITATIONS OF THE MODELS

For CanSPBL, the model for SPBL in Canterbury, the model components of mean top height, basal area and stems per hectare were created with data aged from three to thirty years. Age zero may not be a good starting point from which to project MTH because MTH at that age (just after planting) is probably random rather than a good representation for a site. Basal area at age zero or one was usually not available until some trees reached 1.40 m at age two in Canterbury. It is recommended that the model be re-calibrated every 5 or 10 years by including new PSP measurements.

All plot measurements in SPBL have been taken after the first and only thinning. Thus, different thinning regimes could not be reflected in the model. CanSPBL was mainly for stands on the Canterbury plains and foothills in SPBL's estate. The suitability for coastal sands was not examined.

For CanJuv, the model for juvenile growth in Canterbury, the established model components of mean top height, basal area and survival were expected to be suitable for predictions younger than five years old. The predictive ability beyond five years is unknown. More data covering longer-term measurements and covering wider sites with more detailed silvicultural options are needed to improve the model (Mason and Milne, in prep.).

10.3 SOME SPECIFIC VIEWS

10.3.1 Explanatory variables for modelling: altitude, rainfall and SI

Rainfall had a closer relationship to the juvenile model components of mean top height, basal area, and survival than altitude, which was different to the model CanSPBL. The rationale might be firstly that the data used in the juvenile growth model were so widely distributed that tree growth could not be ascribed to altitude as consistently as rainfall, and secondly that the use of the average rainfall during a trial period instead of a historical average for juvenile growth modelling helped explain growth fluctuations in different years. The coefficient of correlation was 0.9 between altitude and mean rainfall of the last thirty years and it was 0.5 between altitude and mean rainfall during the trial period of juvenile growth. Average rainfall during the trial period was correlated better to growth differences. Altitude was easy to obtain for managers and was also correlated closely with tree growth when the geographical and site condition changes are consistent with altitude. It was the most useful variable explaining variation in tree growth for plains and foothills sites in the mid-Canterbury region of New Zealand.

Site index is commonly used and usually sensitive to site quality, unaffected by thinning from below, and less affected by varying stand density than stand volume. Site index might not be a good indicator (Mason, 1992; Avery and Burkhart, 1994), however, because it often changes periodically with age in terms of climatic fluctuations, and varies with different species on the same site. Height is measured with less accuracy than diameter or basal area per hectare. Despite these limitations, site index was included in two equations in the study: the height-diameter equation at a regional level, and the individual-tree survival equation. Site index was useful in improving the height-diameter model's prediction ability. It contributed a 39.5% reduction in MSE after dbh and age were entered into the model. This variable explained the difference among different stands so well that no other single variable could replace it. For the individual-tree survival model, relative diameter, altitude, initial stocking and projection length were the important predictors. SI was the least important but statistically significant variable.

10.3.2 The relationship between juvenile and whole rotation models

Two models were separately established, namely CanJuv and CanSPBL. Models were created in accordance with modelling objectives. Juvenile growth modelling aimed to predict growth responses to changing environmental and silvicultural conditions. Yield-age equation forms were adopted and two stages of regression were used to identify and incorporate significant variables. Modelling for the whole rotation aimed to project growth for precise use. Use of sigmoid difference equations enabled good estimates to be made given the initial measurements (T_1, Y_1).

The issue of the relationship between juvenile growth and older growth models has arisen since juvenile growth models have been created. Mason *et al.* (1997) displayed the prediction difference between IGM (a juvenile growth model) and PPM88 (a growth model for older crops) with data from one forest as an example. Many theoretical and practical aspects were considered for a smooth transition between models. Linking the two models smoothly may be useful but may be complicated when data sources, model forms and variables included in both models are different.

Necessary conditions for smooth linkages between models are to achieve the same prediction value and the same derivative at linkage points for both juvenile and older curves. The alternative is probably to formulate a compatible piecewise equation and obtain parameter estimates for both juvenile growth and older growth simultaneously during the fitting process. A long-term monitoring and measurement system is required to explore the long-term effects of establishment treatments, and how these can be represented in models of older tree growth. It is not logical to test for predictions of the same changes at their linkage points, given two separate and incompatible data sources.

No attempt was made in the study to link the juvenile growth model with the model for older growth. It was believed that building a whole rotation model by merging the juvenile growth data with older growth data was more meaningful than linking separately established models of juvenile and older growths.

10.4 FUTURE RESEARCH IN MODELLING IMPROVEMENT

Quality models demand quality work at all stages in the modelling procedures. Any violation of statistical and biological principles may introduce bias. No model can be 100% accurate but models built on a sound basis using suitable techniques throughout the whole modelling process can minimise bias. The following aspects are emphasised for further research.

(1) To clarify the issue of data structure in organising interval lengths

Borders *et al.* (1987) stated that a data structure with non-overlapping short intervals was best for one equation form and a data structure with longest intervals was best for the other of two equations they tried. It was concluded that the projection ability with different data structures was model form dependent. Lee (1998) demonstrated all possible mixed intervals as the best data structure in a study for growth and yield in Douglas-fir plantations in the South Island of New Zealand. This study here found that a data structure of all possible intervals led to the best projections of longer intervals, but the benefit was not proved statistically significant in comparison with two other data structures. With these three studies, it might be concluded that the projection ability of different data structures might be data source dependent, as different modellers produced different results using local data sets. None of the three empirical studies have solved the issue universally and fundamentally. The path-invariance property of equations employed may result in the same final yield but in practice when measurement error and annual weather fluctuations are considered, yield increment is more likely to be accurate with longer intervals. Further study may be needed to look at the theoretical aspects with both ideal data and data with various flaws (unbalanced intervals in terms of maximum length on different sites, for example).

(2) Improve the use of two scales of models: stand models and tree models

As described in Chapter 8, there may be three types of relations between a stand model and a tree model: (i) mutually independent, (ii) parallel but compatible and (iii) one dependent on the other. Stand models are generally believed accurate in New Zealand

(Garcia, 1991; Goulding, 1994) and in other countries (Somers and Nepal, 1994). This study adopted the same assumption and disaggregative adjustments were made to individual-tree projections (third type - tree model depends on stand model). The assumption should be justified in future studies, however. This may change as our ability in developing tree models reaches more accurate projections, especially for tree mortality. In the real world, a stand is accurately the aggregation of all individuals and stand volume is simply calculated with summation of volume of individual trees.

CHAPTER 11

SUMMARY OF CONCLUSIONS

This study involved the development of models for whole rotations of radiata pine with sigmoid difference equations, juvenile growth with yield-age equations and height-diameter relationships. Other achievements included database design, inventory sampling evaluation, and model validation.

11.1 ESTABLISHMENT AND ASSESSMENT OF A DATABASE

A new database for permanent sample plot measurements was established using a database management system. Multiple levels of tables and attributes in each table were identified. A relational database was structured and used to manipulate data. Main variables were computed and data for growth modelling were derived from the database.

The main variables for growth modelling were described in both tabular and graphical forms. There were a total of 1200 plots consisting of 2664 repeated plot measurements available for modelling. Stand age ranged from 3 to 31 years. The site Index was from 14 to 26 m at age 20. Basal area reached a maximum of 100 m^2 per hectare. Stocking ranged from 200 to 3000 stems per hectare but the majority lay between 400 to 1000 stems per hectare.

11.2 Modelling of Height-Diameter Relationships

A study was carried out for modelling height-diameter relationships of *Pinus radiata* at stand and regional levels in Canterbury. Sixteen functional forms were evaluated for stands varying considerably in stand age, site index, altitude and the number of trees sampled. A member of the Petterson equation series with the exponent –5.0 and the two-parameter Richards' equation led to the least mean square error overall. The former (Equation 11.1), in particular, has the desired properties of being transformable to a linear form and it was adopted in data manipulation within the study database. Three parameter equations offered limited benefits at huge computational expense for many stands and failed to converge when fitting for a few stands.

$$h = 1.4 + (\alpha + \frac{\beta}{d})^{(-5.0)} \tag{11.1}$$

A height model at a regional scale (Equation 11.2) was obtained by identifying and incorporating the most important variable, stand age, the second most important variable, site index, and the less important but statistically significant variable, altitude, into the selected Petterson family equation. The inclusion of these three variables resulted in a reduction of 72% in the mean square error. It is expected that 90% of predictions should be within ±2.0 m and the model could be useful especially when no samples of height measurements are available.

$$h = 1.40 + [0.695955 + 0.666983T^{-0.5} - 0.106771\ln(SI) + (0.954201 + 0.000741Alt)/d]^{-5} \tag{11.2}$$

11.3 Examination of the Existing Model CANTY

The existing growth and yield model CANTY was examined to determine how closely the model's behaviour fitted the trees in SPBL's estate. Regression analyses and graphical procedures revealed that the model overestimated MTH by 1 m on average. Residuals of MTH, basal area/ha and volume/ha showed significantly close relationships with altitude. The model produced more serious overestimates for lower altitudes on the Canterbury plains and underestimates for higher altitudes on hill sites. No significant bias was detected for stocking projections.

11.4 DEVELOPMENT OF A NEW MODEL, CANSPBL

A new model was built for use at both stand and tree level. The stand model components of mean top height, basal area, stocking, volume and diameter distribution were separately developed by choosing the most suitable equations, using appropriate data structures, and identifying and incorporating significant explanatory variables to improve prediction precision.

A data structure of all possible intervals was chosen in this study for fitting projection equations because it resulted in the least bias for long projections. It seemed that though it could not be proven statistically to be better than all others, neither would it provide a poorer fit.

A polymorphic Schumacher equation displayed the best fit for both MTH and basal area. The altitude effect was formulated and included in the selected equation and it significantly improved the model with a reduction of mean square error by 17% for MTH and 41% for basal area. Altitude represents the difference within and between plains and hills, and may integrate such environmental factors as rainfall, soil, and others. The fitted equation for mean top height and basal area/ha were Equations 11.3 and 11.4, respectively.

$$H_2 = \exp[\ln(H_1)((T_1 + t_0)/(T_2 + t_0))^{\beta} + (\alpha_0 + \alpha_1 Alt + \alpha_2 (Alt - 250)X)/1000 \\ (1 - ((T_1 + t_0)/(T_2 + t_0))^{\beta})] \tag{11.3}$$

$$G_2 = \exp[\ln(G_1)(T_1/T_2)^{\beta} + (\alpha_0 + \alpha_1 Alt + \alpha_2 (Alt - 250)X)/1000(1 - (T_1/T_2)^{\beta})] \tag{11.4}$$

(X=0 when altitude<250, and X=1 when altitude≥250)

A three-parameter difference equation (Equation 11.5) was chosen for stems per hectare. The effect of altitude on live stocking was far less than on basal area and MTH and it was close to but not quite significant at the 5% level. Not much improvement could be obtained with its inclusion into the model.

$$N_2 = (N_1^{c} + a\,(T_2^{b} - T_1^{b}))^{(1/c)} \tag{11.5}$$

A stand volume equation (Equation11.6) displayed a better fit than the traditional form of V/G=α+βH and a reduction of mean square of error by 30% was gained. Future volumes had been estimated with the projected future basal area and future MTH. As a result no apparent bias could be detected.

$$V = \alpha G^{\beta} H^{\gamma} \tag{11.6}$$

Projection equations for maximum diameter and standard deviation were established (Equations 11.7 and 11.8) and a system for producing stand tables with a reverse Weibull function was described. It was shown that a projection of diameter distribution for all individual stands resulted in an unbiased fit within the established system.

$$d_{\max 2} = \exp[\ln(d_{\max 1})(T_1 / T_2)^{\beta} + (\alpha_0 + \alpha_1 Alt)/1000 \ (1 - (T_1 / T_2)^{\beta})] \tag{11.7}$$

$$d_{std 2} = \exp[\ln(d_{std 1})(T_1 / T_2)^{\beta} + (\alpha_0 + \alpha_1 Alt)/1000 \ (1 - (T_1 / T_2)^{\beta})] \tag{11.8}$$

Standard deviation of a stand was estimated with a cluster sampling formulation which accounts for variation within and between plots in a stand. An analysis of the bias of estimating stand variance with the mean of plot variance showed that underestimation was not serious for most stands but serious for a few. On average, the difference between the mean of plot variance and stand variance was a negative value -0.93 cm^2.

Two components of diameter and survival needed for completion of an individual-tree modelling system were built. For the projection of the diameter of individual-trees, an approach based on relative basal area and another based on sigmoid difference equations were compared. Results showed that models based on relative basal area performed much better than models based on sigmoid difference equations. Equation 11.9 led to a best fit with 90% of residuals lying within ±3.1 cm for projections of up to 10.5 years. Information of basal area and stocking at a plot level contributed to the reliability of diameter projections at tree levels. A logistic regression procedure showed that relative diameter, altitude, the interaction between initial stocking and projection interval length, and site index were significantly correlated to the probability of tree death (Equation 11.10). Relative diameter was negatively correlated with tree death, which means the smaller trees in a plot were more likely to die. Trees died more often on poorer sites, higher altitudes, and higher initial stockings with longer projection length.

$$d_2 = \sqrt{\tfrac{1}{0.00007854}\tfrac{G_2}{N_2}[0.00007854d_1^2 /(\tfrac{G_1}{N_1})]^{(T_2/T_1)^\beta}} \quad (11.9)$$

$$P(death) = (\exp(\alpha_0 - \alpha_1 Rd_1 + \alpha_2 Alt + \alpha_3 N_1 \times TimeInc - \alpha_4 SI)^{-1} + 1)^{-1} \quad (11.10)$$

Height can be estimated with diameter using a height-diameter curve (Equation 11.2). Tree volume can be estimated with dbh and height using an existing tree-volume equation which is for Canterbury region.

Two sources of data at a plot and stand level were used to examine the performance of central core components of the new model, CanSPBL. Residual analyses by means of a regression test and graphical plotting indicated that the model showed no significant bias for MTH, basal area per hectare, stems per hectare and volume per hectare when used at both a plot and stand level in SPBL's estate.

A computer simulation program was developed for use at both stand and tree level. It produces estimates for multiple stands or plots simultaneously, given recent measurements. It is simple and flexible but applicable only in an Excel 97 environment.

11.5 Modelling of Juvenile Growth

Nine sets of experimental data and twelve temporary plots were used to build the juvenile growth model components of mean top height, sectional area at ground level, and survival. Thirty-one plots were used to build a model of basal area. Stepwise linear regression showed that mean annual rainfall during trial period and weeding were significant factors determining the growth of mean top height and basal area. Higher rainfall could reduce mortality significantly but affected survival less than mean top height and basal area. Annual rainfall and weeding described growth differences within and among experiments.

The four equations for mean top height, sectional area at ground level, survival and basal area were Equations 11.11 to 11.14, respectively.

$$H_T = H_0 + e^{(a_0 + a_1 RAINFALL)} T^{(b_0 + b_1 WEEDING)} \tag{11.11}$$

$$G_{GL.T} = G_{GL.0} + e^{a} N_0 T^{(b_0 + b_1 RAINFALL + b_2 WEEDING)} \tag{11.12}$$

$$S_T = e^{(a_0 + a_1 RAINFALL)} T^{b} \tag{11.13}$$

$$G_T = \alpha G_{GL.T} [1 - \exp(-\beta G^{\gamma}{}_{GL.T})] \tag{11.14}$$

ACKNOWLEDGEMENTS

My acknowledgement first goes to my supervisor, Dr. Euan G. Mason, School of Forestry, University of Canterbury. He gave guidance, advice, and assistance in searching for data, and solving various problems throughout the research. It was his support that made the thesis possible.

A favourable environment for the study was given by the School of Forestry (SOF). The staff of SOF have been excellent and my fellow postgraduates were also very supportive. In particular the help of SOF in employing a typist to enter PSP data for the study is appreciated.

The awards of the TW Adams Scholarship during 1995 and the University of Canterbury Doctoral Scholarship during 1998 are acknowledged. They gave important financial and moral support to the author of this thesis.

Selwyn Plantation Board Limited made its PSP data available for the study. Thanks are expressed to Mr W. P. Studholme and Mr H. W. Stevenson for their support in providing all needed information and efforts in obtaining data from other companies.

I am grateful to Jane Thomas, Pasefika Education and Employment Training Organisation, Christchurch, for her work in proof reading this thesis. The valuable corrections given by examiners, Professor H. E. Burkhart and Dr. A.G.D. Whyte, are also deeply appreciated.

I would like to express my gratitude to my parents for their close attention to my early education. My wife, Changxiang, and my son, Wenda, contributed love and support.

My thanks to everyone who helped in this study.

REFERENCES

Arabatzis, A.A. and Burkhart, H.E., 1992, An evaluation of sampling methods and model forms for estimating height-diameter relationships in loblolly pine plantations, Forest Science, 38(1): 192-198

Arnold, D.R., 1994, PlotInfo: A system to store, manipulate, and retrieve forest inventory data through the Geographic Information System MapInfo, Forestry Dissertation for Bachelor, School of Forestry, University of Canterbury, 59 pages

Avery, T.E. and Burkhart, H.E., 1994, Forest measurements, 4th Edition, McGraw-Hill, New York, 408 p.

Avila, O.B. and Burkhart, H.E., 1992, Modelling survival of loblolly pine trees in thinned and unthinned plantations, Canadian Journal of Forest Research, 22:1878-1882

Bailey, R.L. and Dell, T.R., 1973, Quantifying diameter distributions with the Weibull function, Forest Science, 19(2): 97-104

Ballard, R., 1978a, Effect of slash and soil removal on the productivity of second rotation radiata pine on a pumice soil, New Zealand Journal of Forestry Science, 8(2): 248-258

Ballard, R., 1978b, Use of fertilisers at establishment of exotic forest plantations in New Zealand, New Zealand Journal of Forestry Science, 8(1): 70-104

Ballard, R. and Will, G.M., 1981, Removal of logging waste, thinning debris, and litter from a *Pinus radiata* pumice soil site, New Zealand Journal of Forestry Science, 18(2): 152-163

Balneaves, J., Christie, M. and Balneaves, J.M., 1988, Long-term growth response of radiata pine to herbaceous weed control at establishment, New Zealand Forestry, 33(3): 24-25

Balneaves, J.M., 1982, Grass control for radiata pine establishment on droughty sites, New Zealand Journal of Forestry, 27(2): 259-276

Bates, D.M. and Watts, D.G., 1980, Relative curvature measures of nonlinearity, Journal of Royal Statistical Society, 42:1-16

Beekhuis, J., 1966, Prediction of yield and increment in *Pinus radiata* stand in New Zealand, New Zealand Forest Service, Forest Research Institute Technical Paper No.49, 39 pages

Belli, K.L., 1987, FPES: A framework for modelling the artificial regeneration system, In: Ek, A.R., Shifley, S.R. and Burk, T.E. (editors) Proceedings of the IUFRO symposium on Forest growth modelling and prediction, Minneapolis, Minnesota, USDA Forest Service, General Technical Report NC-120, 361-368

Belli, K.L. and Ek, A.R., 1988, Growth and survival modelling for planted conifers in the Great Lakes region, Forest Science, 34(2): 458-473

Biging, G.S. and Dobbertin, M., 1995, Evaluation of competition indices in individual tree growth models, Forest Science, 41(2): 360-377

Bitterlich, W., 1947, Die Winkelzahlmessung (Measurement of basal area per hectare by means of angle measurement), Allgm. Forst- u. Hotzwirts. Ztg. 58: 94-96, Forestry Abstracts 10: 1473

Bitterlich, W., 1948, Die Winkelzahlprobe (The angle-count sample plot), Allgm. Forst- u. Holzwirts. Ztg. 59: 4-5, Forestry Abstracts 10: 2314

Bliss, C.I. and Reinker, K.A., 1964, A lognormal approach to diameter distributions in everaged stands, Forest Science, 10(3): 350-360

Borders, B.E., Bailey, R.L. and Clutter, M.L., 1987a, Forest growth models: parameter estimation using real growth series, In: Ek, S.R., Shifley, S.R. and Burk, T.E.

(editors) Forest Growth Modelling and Prediction, Proceedings IUFRO Conference, Minneapolis, MN, USDA forest Sèrvice, General Technical Report NC-120, p. 660-667

Borders, B.E. and Patterson, W.D., 1990, Projecting stand tables: a comparison of the Weibull diameter distribution method, a percentile-based projection method, and a basal area growth projection method, Forest Science, 36(2): 413-424

Borders, B.E., Souter, R.A., Bailey, R.L. and Ware, K.D., 1987b, Percentile-based distributions characterize forest stand tables, Forest Science, 33(2): 570-576

Bossel, H., 1991, Modelling forest dynamics: moving from description to explanation, Forest Ecology and Management, 42:129-142

Botkin, D.B., 1993, Forest Dynamics: An Ecological Model, Oxford University Press, 399 pages

Brackett, M., 1973, Notes of tariff tree volume computation, Washington, Department of National Research, Research Note SE-179

Bruce, D., 1926, A method of preparing timber yield tables, Journal of Agricultural Research, 32:543-557

Bruce, D. and Wensel, L.C., 1987, Modelling forest growth: Approaches, definitions, and problems, In: Ek, A.R., Shifley, S.R. and Burk, T.E. (editors) Proceedings of the IUFRO symposium on forest growth modelling and prediction, Minneapolis, Minnesota, USDA, Forest Service General Technical Report NC-120, 1-8

Buford, M.A., 1986, Height-diameter relationships at age 15 in loblolly pine seed sources, Forest Science, 32(3): 812-818

Bunnell, F.L., 1989, Alchemy and uncertainty: what good are models?, USDA Forest Service, General Technical Report, PNW-GTR-232, 27 pages

Burkhart, H.E., 1977, Stand modelling for radiata pine in New Zealand, New Zealand Journal of Forestry, 22 (2): 297-307

Burkhart, H.E. and Tennent, R.B., 1977, Site index equations for radiata pine in New Zealand, New Zealand Journal of Forestry Science, 7(3): 408-416

Cao, Q.V. and Burkhart, H.E., 1984, A segmented distribution approach for modeling diameter frequency data, Forest Science, 30(1): 129-137

Chisman, H.H. and Schumacher, F.X., 1940, On the tree area ratio and certain of its applications, Journal of Forestry, 26(1): 55-69

Clifton, N.C., 1969, Resin pockets in Canterbury Radiata Pine, New Zealand Journal of Forestry, 14(1): 38-49

Clutter, J.L., 1963, Compatible growth and yield models for loblolly pine, Forest Science, 9(3): 354-371

Clutter, J.L. and Allison, B.J., 1974, A growth and yield model for *Pinus radiata* in New Zealand, In: Fries, J. (editor) Growth Models for Tree and Stand Simulation, Proceedings of IUFRO Working Party S4.01-4 meetings, Dep. For. Yield Res., Royal Coll. For., Stockholm , p.136-160

Clutter, J.L., Fortson, J.C., Pienaar, L.V., Brister, G.H. and Bailey, R.L., 1983, Timber Management: A Quantitative Approach, New York, John Wiley and Sons, 333 pages

Clutter, J.L. and Jones, E.P., 1980, Prediction of growth after thinning in old-field slash pine plantations, USDA Forest Service, Research Paper, SE-217

Codd, E.F., 1970, A relational model for large shared data banks, Communications of the ACM, 13(6): 377-387

Codd, E.F., 1990, The relational model for database management, Version 2, Reading, MA: Addison Wesley, 538 pages

Connolly, T.M., Begg, C.E. and Strachan, A.D., 1996, Database Systems: A Practical Approach to Design, Implementation and Management, England, Addison-Wesley, 839 pages

Curtis, R.O., 1967, Height-diameter and height-diameter-age equations for second-growth Douglas-fir, Forest Science, 13:365-375

Daniels, R.F. and Burkhart, H.E., 1975, Simulation of individual tree growth and stand development in managed loblolly pine plantations, Div. of For. and Wildlife Res., Va. Polytechnic Inst. and State Univ., FWS-5-75

Daniels, R.F. and Burkhart, H.E., 1988, An integrated system of forest stand models, Forest Ecology and Management, 23:159-177

Dolph, L.D., 1989, Height-diameter equations for young-growth Red-fir in California and Southern Oregon, US Pacific Southwest Research Station, Research Note PSW-408

Dolph, L.D., Mori, S.R. and Oliver, W.W., 1995, Height-diameter relationships for conifer species on the Blacks Mountain experimental forest, US Pacific Southwest Research Station, Research Note PSW-RN-418

Dyck, W.J. and Bow, C.A., 1992, Environmental impacts of harvesting, Biomass and Bioenergy, 2(1-6): 173-191

Dyck, W.J., Mees, C.A. and Comerford, N.B., 1989, Medium-term effects of mechanical site preparation on radiata pine productivity in New Zealand - a retrospective approach, In: Dyck, W.J. and Mees, C.A. (editors) Proceedings of the IEA/BE A3 workshop on 'Research strategies for long-term site productivity', Seattle, WA, 1988, New Zealand Forest Research Institute Bulletin No. 152, 79-92

Ek, A.R., 1973, Performance of regression models for tree height estimation with small sample sizes, in 'Statistics in Forest Research', In: Proceedings of 4th IUFRO Conference of Advisory Group of Forest Statisticians , 20-24 August 1973, Vancouver, B.C. Canada, 67-80

Ek, A.R., Issos, J.N. and Bailey, R.L., 1975, Solving for Weibull diameter distribution parameters to obtain specified mean diameters, Forest Science, 21(3): 290-292

Eyles, G.O., 1986, *Pinus radiata* site index rankings for New Zealand, New Zealand

Forestry, 31(2): 19-22

Farrell, P.W., 1984, Radiata pine residue management and its implications for site productivity on sandy soils, Australian Forestry, 47(2): 95-102

Flewelling, J.W. and De Jong, R., 1994, Considerations in simultaneous curve fitting for repeated height-diameter measurements, Canadian Journal of Forest Research, 24(7): 1408-1414

Freese, F., 1960, Testing accuracy, Forest Science, 6(2): 139-145

Freese, F., 1962, Elementary forest sampling, USDA Forest Service, Agriculture Handbook 232, 91 pages

Gadd, A. and Wold, H., 1964, The Janus coefficient: a measure for the accuracy of prediction, In: Wold, H. O. A., (Editor), Econometric Model Building: Essays on the causal chain approach, North-Holland, Amsterdam, 229-235

Garcia, O., 1974, Height-diameter equations for *Pinus radiata*, Instituto Forestal, Chile, Nota tecnica, No. 19

Garcia, O., 1981, Simplified methods-of-moments estimation for the Weibull distribution, New Zealand Journal of Forestry Science, 11:304-307

Garcia, O., 1983, A stochastic differential equation model for the height growth of forest stands, Biometrics, 39(4): 1059-1072

Garcia, O., 1984, New class of growth models for even-aged stands: *Pinus radiata* in Golden Downs Forest, New Zealand Journal of Forestry Science, 14(1): 65-88

Garcia, O., 1988, Growth modelling - a (re)view, New Zealand Forestry, 33(3): 14-17

Garcia, O., 1991, What is a diameter distribution?, In: Proceedings of the IUFRO symposium on "Integrated Management Information System", Tsukuba, Japan, October 13-16, 1991, 19 pages

Garcia, O., 1994, The state-space approach in growth modelling, Canadian Journal of Forest Research, 24(9): 1894-1903

Garman, S.L., Acker, S.A., Ohmann, J.L. and A., S.T., 1995, Asymptotic height-diameter equations for twenty-four species in Western Oregon, Forest Research Laboratory, Oregon State University, Research contribution 10

Genetics and Tree Improvement Research Field, 1987, Which radiata pine seed should you use?, New Zealand Forest Research Institute What's New In Forest Research No. 157, 4 pages

Glover, G.R. and Hool, J.N., 1979, A basal area ratio predictor of loblolly pine plantation mortality, Forest Science, 25:275-282

Goulding, C.J., 1979, Validation of growth models used in forest management, New Zealand Journal of Forestry, 24(1): 108-124

Goulding, C.J., 1994, Development of growth models for Pinus radiata in New Zealand - experience with management and process models, Forest Ecology and Management, 69:331-343

Goulding, C.J., 1995, Measurement of trees; Growth and yield models, In D. Hammond (editor): 'NZIF 1995 Forestry Handbook', 1995, New Zealand Institute of Forestry (Inc.), Christchurch, p.104-107 and p.111-114

Goulding, C.J. and Lawrence, M.E., 1992, Inventory practice for managed forests, FRI Bulletin, No. 171

Greacen, E.L. and Sands, R., 1980, Compaction of forest soils: a review, Australia Journal of Soil Research, 18:163-189

Gregoire, T.G. and Reynolds, M.R., 1988, Accuracy testing and estimation alternatives, Forest Science, 34(2): 302-320

Hafley, W.L. and Schreuder, H.T., 1977, Statistical distributions for fitting diameter and height data in even-aged stands, Canadian Journal of Forest Research, 7(3): 481-487

Hägglund, B., 1981, Forecasting growth and yield in established forests: an outline and analysis of the outcome of a subprogram within the HUGIN project, Swedish

University of Agriculture Science, Dep. For. Survey, Report 31, 145 pages

Hall, M., 1985, Establishment of radiata pine on a high altitude second rotation site, II: Effect of site preparation on early survival and growth, Australian Forestry, 48(2): 79-83

Hamilton, D.A., Jr., 1990, Extending the range of applicability of an individual tree mortality model, Canadian Journal of Forest Research, 20(8): 1212-1218

Hamilton, D.A. and Edwards, B.M., 1976, Modelling the probability of individual tree mortality, USDA Forest Service Research Paper, Intermountain Forest and Range Experiment Station, INT-185

Hann, D.W., 1980, Development and evaluation of an even- and uneven-aged ponderosa pine/Arizone fescue stand simulator, USDA Forest Service Research Paper, INT-267, 95 pages

Henriksen, H.A., 1950, Height-diameter curve with logarithmic diameter: brief report on a more reliable method of height determination from height curves, introduced by the State Forest Research Branch, Dansk skovforen tidsskr, 35:194-202

Hetherington, M.W. and Balneaves, J.M., 1973, Ripping in tussock country improves Radiata growth, Forest Industries Review, 4(12): 2-7

Hinds, H.V. and Reid, J.S., 1957, Forest trees and timbers of News Zealand, New Zealand Forest Service, Bulletin No. 12

Huang, S.M., Titus, S.J. and Wiens, D.P., 1992, Comparison of nonlinear height-diameter functions for major Alberta tree species, Canadian Journal of Forest Research, 22(9): 1297-1304

Hunter, I.R., 1996, The occurrence and treatment of magnesium deficiency in radiata pine in New Zealand, New Zealand Forest Research Institute Bulletin No. 172, 136 pages

Hunter, I.R. and Gibson, A.R., 1984, Predicting *Pinus radiata* site index variables from environmental variables, New Zealand Journal of Forestry Science, 14 (1): 53-64

Hunter, I.R. and Graham, J.D., 1982, Growth response of phosphorus-deficient *Pinus radiata* to various rates of superphosphate fertiliser, New Zealand Journal of Forestry Science, 12(1): 49-61

Hunter, I.R. and Hoy, G.F., 1983, Growth and nutrition of *Pinus radiata* on a recent coastal sand as affected by nitrogen fertiliser, New Zealand Journal of Forestry Science, 13(1): 3-13

Hunter, I.R., Rodgers, B.E., Duningham, A., Prince, J.M. and Thorn, A.J., 1991, An atlas of radiata pine nutrition in New Zealand, New Zealand Forest Research Institute Bulletin No. 165, 24 pages

Husch, B., 1971, Planning a forest inventory, FAO, Italy, 120 pages

Husch, B., Miller, C. and Beers, T.W., 1982, Forest mensuration, New York, John Wiley & Sons Inc., 402 pages

Hyink, D.M. and Moser, J.W., 1983, A generalised framework for projecting forest yield and stand structure using diameter distributions, Forest Science, 29(1): 85-95

Jackson, D.S. and Gifford, H.H., 1974, Environmental variables influencing the increment of radiata pine: (1) periodic volume increment, New Zealand Journal of Forestry Science, 4:3-26

Jackson, D.S., Gifford, H.H. and Chittenden, J., 1976, Environmental variables influencing the increment of radiata pine: (2) Effects of seasonal drought on height and diameter increment, New Zealand Journal of Forestry Science, 5:265-286

Jacobs, M.R., 1937, The detection of annual stages of growth in the crown of *Pinus radiata* Commonwealth Forestry Bureau, Canberra Bulletin, No.10, 19 pages

James, R.N., 1976, Implications for silviculture from the Tarawera valley regimes trial, New Zealand Journal of Forestry Science, 6(2): 171-181

Johnson, N.L., 1949a, Bivariate distributions based on simple translation systems, Biometrika, 36:297-304

Johnson, N.L., 1949b, Systems of frequency curves generated by methods of translation, Biometrika, 36:149-176

Knoebel, B.R. and Burkhart, H.E., 1991, A bivariate distribution approach to modelling forest diameter distributions at two points in time, Biometrics, 47:241-253

Knowe, S.A., 1994, Effect of competition control treatments on height-age and height-diameter relationships in young Douglas-fir plantations, Forest Ecology and Management, 67(1-3): 101-111

Knowe, S.A., Ahrens, G.R. and DeBell, D.S., 1997, Comparison of diameter-distribution-prediction, stand-table-projection, and individual-tree-growth modelling approaches for young red alder plantations, Forest Ecology and Management, 98(1): 49-60

Koesmarno, H.K., 1996, Class-size percentile transformation for reconstruction a distribution function, Journal of Applied Statistics, 23(4): 447-458

Krajicek, J.E., Brinkman, K. A. and Gingrich, S.F., 1961, Crown competition - a measure of density, Forest Science, 7:35-42

Kuru, G.A., Whyte, A.G.D. and Woollons, R.C., 1992, Utility of reverse Weibull and extreme value density functions to refine diameter distribution growth estimates, Forest Ecology and Management, 48(1-2): 165-174

Lane Poole, C.E., 1944, Monterey pine: Space and increment, Australian Forestry, 8(1): 20-29

Lange, P.W., de Ronde, C. and Bredenkamp, B.V., 1987, The effects of different intersities of pruning on the growth of *Pinus radiata* in South Africa, South African Forestry Journal, No. 143:30-36

Lappi, J., 1997, A longitudinal analysis of height/diameter curves, Forest Science, 43(4): 555-570

Larsen, D.R. and Hann, D.W., 1987, Height-diameter equations for seventeen tree species in Southwest Oregon, Research Paper Forest Research Laboratory, School of Forestry, Oregon State University, 49): 16 pages

Latif, M.A., Rahman, M.B. and Das, S., 1996, Initial growth performance for *Pinus caribaea* in Bangladesh, In: Skovsgaard, J.P. and Johannsen, V.K. (editors) The IUFRO Conference on 'Modelling Regeneration Success and Early Growth of Forest Stands', June, 1996, Copenhagen, Denmark, Ministry of environment and energy, Danish forest and landscape research institute, p. 196-199

Leary, R.A., 1979, Design, In: A generalised forest growth projection system applied to the lake states region, USDA Forest Service, General Technical Report NC-49, p.5-15

Leary, R.A., Nimerfro, K., Holdaway, M., Brand, G., Kolka, R., Wolf, A. and Burk, T., 1996, The importance of initial height growth in quaking aspen height growth models expressed as 2nd order differential equations, In: Skovsgaard, J.P. and Johannsen, V.K. (editors) The IUFRO Conference on 'Modelling Regeneration Success and Early Growth of Forest Stands', June, 1996, Copenhagen, Denmark, Ministry of environment and energy, Danish forest and landscape research institute, p. 200-210

Ledgard, N.J. and Belton, M.C., 1985, Exotic trees in the Canterbury high country, New Zealand Journal of Forestry Science, 15(3): 298-323

Lee, S. H., 1998, Modelling growth and yield of Douglas-fir using different interval lengths in the South Island of New Zealand, PhD Thesis, University of Canterbury, Christchurch, New Zealand, 239 pages,

Lenhart, J.D. and Clutter, J.L., 1971, Cubic foot yield tables for old field loblolly pine plantations in the Georgia Piedmont, Ga. For. Res. Council, Report 22(3)

Lewis, E.R., 1954, Yield of unthinned *Pinus radiata* in New Zealand, New Zealand Forest Service, Forestry Research Institute, Forest Research Notes, 1(10): 19 pages

Lindsay, S.R., Wood, G.R. and Woollons, R.C., 1996, Stand table modelling through the Weibull distribution and usage of skewness information, Forest Ecology and Management, 81:19-23

Liu, C.M., Leuschner, W.A. and Burkhart, H.E., 1989, A production function analysis of loblolly pine yield equations, Forest Science, 35(3): 775-788

Loague, K. and Green, R.E., 1991, Statistical and graphical methods for evaluation solute transport models: Overview and application, Journal Contam. Hydrol., 7:51-73

Maclaren, J.P., 1993, Radiata pine growers' manual, New Zealand Forest Research Institute Bulletin, No. 184

Madgwick, H.A.I., 1994, *Pinus radiata*: biomass, form and growth, Rotorua, New Zealand, 428 pages

Maltamo, M., Puumalainen, J. and Paivinen, R., 1995, Comparison of beta and Weibull distribution for modelling basal area diameter distribution in stands of Pinus sylvestris and Picea abies, Scandinavian Journal of Forest Research, 10(3): 284-295

Mason, E.G., 1989, Toppling in New Zealand---where do we go from here? In A. Somerville, S. Wakelin, and L. Whitehouse (editors), Workshop on wind damage in New Zealand exotic forests, FRI Bulletin 146

Mason, E.G., 1992, Decision-support systems for establishing radiata pine plantations in the Central North Island of New Zealand, PhD Thesis, University of Canterbury, Christchurch, New Zealand, 243 pages

Mason, E.G. and Cullen, A.W.J., 1986, Machinery for forest establishment and silviculture, IN: Levack, H. (Ed.), "Forestry Handbook", New Zealand Institute of Foresters, 60-62

Mason, E.G., Cullen, A.W.J. and Rijkse, W.C., 1989, Growth of two Pinus radiata stock types on ripped and ripped/bedded plots at Karioi forest, New Zealand Journal of Forestry Science, 18(3): 287-296

Mason, E.G. and Milne, P.G., 1999, Effects of weed control, fertilisation and soil cultivation on the growth of *Pinus radiata* D.Don at mid-rotation in Canterbury, New Zealand, in prep.

Mason, E.G., South, D.B. and Zhao, W., 1995, Interactions among seedling grade, weed control, and soil cultivation for *Pinus radiata* in the Central North Island of New Zealand, In: Gaskin, R.E., & J.A. Zabkeiwicz (Eds), Popular summaries from 2nd International conference on forest vegetation management, Forest Research Institute Bulletin 192: 107-109

Mason, E.G., South, D.B. and Zhao, W., 1996, Performance of *Pinus radiata* in relation to seedling grade, weed control, and soil cultivation in the Central North Island of New Zealand, New Zealand Journal of Forestry Science, 26 (1/2):173-183

Mason, E.G. and Whyte, A.G.D., 1997, Modelling initial survival and growth of radiata pine in New Zealand, Acta Forestalia Fennica 255, 38 pages

Mason, E.G., Whyte, A.G.D., Woollons, R.C. and Richardson, B., 1997, A model of the growth of juvenile radiata pine in the Central North Island of New Zealand: links with older models and rotation-length analyses of the effects of site preparation, Forest Ecology and Management, 97(2): 187-195

Mayer, D.G. and Butler, D.G., 1993, Statistical validation, Ecological Modelling, 68:21-32

McEwen, A.D., 1978, N. Z. Forest Service computer system for permanent sample plots, In FRI Symposium - 'Mensuration for Management Planning of Exotic Forest Plantations', Symposium No. 20, p.235-252

Mead, D.J. and Gadgil, R.L., 1978, Fertiliser use in established radiata pine stands in New Zealand, New Zealand Journal of Forestry Science, 8(1): 105-134

Menzies, M.I., 1988, Seedling quality and seedling specifications for radiata pine, New Zealand Forest Research Institute, What's New in Forest Research, No. 171

Menzies, M.I. and Chavasse, C.G.R., 1982, Establishment trials on frost-prone sites, New Zealand Journal of Forestry, 27(1): 33-49

Ministry of Forestry, 1994, Regional Studies Canterbury, Wellington, New Zealand, 45 pages

Moeur, M. and Ek, A.R., 1981, Plot, stand, and cover-type aggregation effects on projections with an individual tree based stand growth model, Canadian Journal of Forest Research, 11(2): 309-315

Munro, D.D., 1973, Forest growth models - a prognosis, IN: Fries, J. (editor) Growth models for tree and stand simulation, Royal College of Forestry, Research Note 30, Stockholm, p. 7-12

Murphy, G., 1983, *Pinus radiata* survival, growth and form four years after planting off and on skidtrails, New Zealand Journal of Forestry, 28(2): 185-193

Myers, C.A., 1966, Height-diameter curves for tree species subject to stagnation, USDA Forest Servie, Rocky Mountain Forest and Range Experiment Station, Research Note RM-69

Myers, H.A., 1940, A mathematical expression for height curve, Journal of Forestry, 38:415-420

Nelson, T.C., 1964, Diameter distribution and growth of Loblolly pine, Forest Science, 10(1): 105-115

Neter, J. and Wasserman, W., 1974, Applied linear statistical models: regression, analysis of variance, and experimental designs, Georgetown, Ontario, Richard D. IRWIN, INC., 842 pages

New Zealand Forest Research Institute, 1991, FFCalc, Forestry Function Calculator user manual and tutorial, Version 1, NZFRI, Rotorua, New Zealand, FRI software series 15

Nokoe, S., 1978, Demonstrating the flexibility of the Gompertz function as a yield model using mature species data, The Commonwealth Forestry Review 57: 35-42

Nystrom, K. and Kexi, M., 1997, Individual tree basal area growth models for young stands of Norway spruce in Sweden, Forest Ecology and Management, 97(2): 173-185

Pienaar, L.V. and Harrison, W.M., 1988, A stand table projection approach to yield prediction in unthinned even-aged stands, Forest Science, 34(3): 804-808

Pienaar, L.V. and Turnbull, K.J., 1973, The Chapman-Richards generalization of von Bertalanffy's growth model for basal area growth and yield in even-aged stands, Forest Science, 19(1): 2-22

Power, M., 1993, The predictive validation of ecological and environmental models, Ecological Modelling, 68:33-50

Reineke, L.H., 1933, Perfecting a stand density index for even-aged stands, Journal of Agriculture Research, 46:627-638

Reynolds, M.R., 1984, Estimating the error in model predictions, Forest Science, 30(2): 454-469

Reynolds, M.R., Burkhart, H.E. and Daniels, R.F., 1981, Procedures for statistical validation of stochastic simulation models, Forest Science, 27(2): 349-364

Richards, F.J., 1959, A flexible growth function for empirical use, Journal of Experimental Botany, 10:290-300

Ritchie, M.W. and Hann, D.W., 1997, Evaluation of individual-tree and disaggregative prediction methods for Douglas-fir stands in western Oregon, Canadian Journal of Forest Research, 27(2): 207-216

Sands, R., 1983, Physical changes to sandy soils planted to radiata pine, In: Ballard, R. and Gessel, S.P. (editors) IUFRO Symposium on Forest Site and Continuous Productivity, USDA Forest Service, Pacific Northwest Forest and Range Experiment Station, General Technical Report, No. 163, p. 146-152

Sands, R. and Nambiar, E.K.S., 1984, Water relations of Pinus radiata in competition with weeds, Canadian Journal of Forest Research, 14(2): 233-237

SAS Institute Inc., 1990, SAS/STAT user's guide, version 6, Cary, NC

Schaeffer, D.L., 1980, A model evaluation methodology applicable to environmental assessment models, Ecological Modelling, 8:275-295

Schmidt, A.v., 1967, Der rechnerische ausgleich von bestandeshohenkurven, Forstwissenschaftliches zentrat blatt, 86:370-382

Schumacher, F.X., 1939, A new growth curve and its application to timber yield studies, Journal of Forestry, 37: 819-820

Shepherd, K.R., 1961, The effects of low pruning on increment of radiata pine plantations, Forestry Commission of New Sough Wales, Research Note, No. 6:18

Shepherd, K.R., 1967, Influence of low pruning on first thinning in radiata pine plantations, Australian Forestry, 31(1): 45-49

Shepherd, R.W., 1990, Early importation of *Pinus radiata* to New Zealand and distribution in Canterbury to 1885: implications for the genetic makeup of *Pinus radiata* stocks, Part I, Horticulture in New Zealand, 1(1): 33-38

Shiver, B.D. and Borders, B.E., 1996, Sampling Techniques for Forest Resource Inventory, USA, John Wiley & Sons Inc., 356 pages

Shula, R.G., Stand growth model with P fertiliser effects for radiata pine on clay soil, Ministry of Forestry FRI Bulletin, No. 148

Smith, J.L. and Burkhart, H.E., 1984, A simulation study assessing the effect of sampling for predictor variable values on estimates of yield, Canadian Journal of Forest Research, 14:326-330

Snee, R.D., 1977, Validation of regression models: methods and example, Technometrics, 19:415-428

Snowdon, P. and Waring, H.D., 1984, Long-term nature of growth responses obtained to fertiliser and weed control applied at planting and their consequences for forest management, IN: Proceedings of the IUFRO symposium on site and site productivity of fast growing plantations, Pretoria and Petermaritzberg, South Africa, p. 701-711

Soares, P., Tome, M., Skovsgaard, J.P. and Vanclay, J.K., 1995, Evaluating a growth model for forest management using continuous forest inventory data, Forest Ecology and Management, 71(3): 251-265

Somers, G.L. and Nepal, S.K., 1994, Linking individual-tree and stand-level growth models, Forest Ecology and Management, 69(1994): 233-243

SPBL, 1998, Annual Report, Selwyn Plantation Board Limited, Christchurch, New Zealand, 40 pages

Spurr, S.H., 1952, Forest Inventory, New York, The Ronald Press Company, 476 pages

Studholme, B., 1989, Windthrow on the Canterbury plains, In A. Somerville, S. Wakelin, and L. Whitehouse (editors), Workshop on wind damage in New Zealand exotic forests, FRI Bulletin 146, p.28-31

Sullivan, A.D. and Clutter, J.L., 1972, A simultaneous growth and yield model for Loblolly Pine, Forest Science, 18(1): 76-86

Sutton, W.R.J. and Crowe, J.B., 1975, Selective pruning of radiata pine, New Zealand Journal of Forestry Science, 5(2): 171-195

Temu, M.J., 1992, Forecasting yield of Douglas fir in the South Island of New Zealand, PhD Thesis, University of Canterbury, Christchurch, New Zealand, 246 pages

Tennent, R.B., 1982, Individual-tree growth model for Pinus radiata, New Zealand Journal of Forestry Science, 12(1): 62-70

Tennent, R.B., 1986, Intra-annual growth of young Pinus radiata in New Zealand, New Zealand Journal of Forestry Science, 16(2): 166-175

Trorey, L.G., 1932, A mathematical method for the construction of diameter height curves based on site, The Forestry Chronicle, 8:121-132

Turner, R., 1989, Canterbury plains and Lake Taupo forests, In A. Somerville, S. Wakelin, and L. Whitehouse (editors), Workshop on wind damage in New Zealand exotic forests, FRI Bulletin 146

Turvey, N.D. and Cameron, J.N., 1986, Site preparation for a second rotation of radiata pine: soil and foliage chemistry, and effect on tree growth, Australian Forest Research, 16:9-19

Van Larr, A., 1973, Needle - biomass, growth and growth distribution of *Pinus radiata* in South Africa in relation to pruning and thinning, Forstliche Forschungsanstalt Munchen, Nr. 9:283

Vanclay, J.K., 1991, Mortality functions for north Queensland rainforests, Journal of Tropical Forest Science, 4:15-36

Vanclay, J.K., 1994, Modelling forest growth and yield: applications to mixed tropical forests, Wallingford, CAB International, 312 pages

Vanclay, J.K., Skovsgaard, J.P. and Hansen, C.P., 1995, Assessing the quality of permanent sample plot databases for growth modelling in forest plantations, Forest Ecology and Management, 71(1995): 177-186

Vincent, T.G. and Dunstan, J.S., 1989, Register of commercial seedlots issued by the New Zealand Forest Service, Ministry of Forestry, Forest Research Institute Bulletin 144

Von Bertalanffy, L., 1949, Problems of organic growth, Nature, 163(4135): 156-159

Wang, C.H. and Hann, D.W., 1988, Height-diameter equations for sixteen tree species in the central western Willamette valley of Oregon, Oregon State University, Forestry research laboratory, Research Paper 51

Weisberg, S., 1985, Applied linear regression, Wiley, N. Y., 324 pages

West, G.G., Eggleston, N.T. and McLanachan, J.R., 1987, Further development and validation of the EARLY growth model, New Zealand Forest Research Institute Bulletin No.129, 32 pages

West, G.G., Knowles, R.L. and Koehler, A.R., 1982, Model to predict the effects of pruning and early thinning on the growth of radiata pine, NZ FRI Bulletin No. 5, 35 pages

West, P.W., 1995, Application of regression analysis to inventory data with measurements on successive occasions, Forest Ecology and Management, 71:227-234

West, P.W., Ratkowsky, D.A. and Davis, A.W., 1984, Problems of hypothesis testing of regressions with multiple measurements from individual sampling units, Forest Ecology and Management, 7(3): 207-224

Whiteside, I.D., 1990, STANDPAK modelling system for radiata pine, In: James, R.N. and Tarlton, G., L. (editors) Proceedings of the IUFRO symposium on "New approaches to spacing and thinning in plantation forestry", Forest Research Institute Bulletin No. 151, p.106-111

Whyte, A.G.D. and Woollons, R.C., 1992, Diameter distribution growth and yield modelling: recent reviews and perspectives, In: Wood, G.B. and Turner, B. (editors) Proceedings of IUFRO S4.01 & S4.02 Conference on "Integrating forest information over space and time", Canberra

Whyte, A.G.D. and Woollons, R.C., 1990, Modelling stand growth of radiata pine thinned to varying densities, Canadian Journal of Forest Research, 20(7): 1069-1076

Will, G.M., 1985, Nutrient deficiencies and fertiliser use in New Zealand exotic forests, New Zealand Forest Service, Forest Research Institute Bulletin No. 97, 53 pages

Williamson, M.J., 1985, Cultivation in Northland, New Zealand Journal of Forestry, 30(2): 218-231

Woollons, R.C., 1998, Even-aged stand mortality estimation through a two-step regression process, Forest Ecology and Management, 105:189-195

Woollons, R.C. and Hayward, W.J., 1985, Revision of a growth and yield model for radiata pine in New Zealand, Forest Ecology and Management, 11(3): 191-202

Woollons, R.C., Snowdon, P. and Mitchell, N.D., 1997, Augmenting empirical stand projection equations with edaphic and climatic variables, Forest Ecology and

Management, 98:267-275

Woollons, R.C., Whyte, A.G.D. and Xu, L., 1990, The Hossfeld function: an alternative model for depiction stand growth and yield and growth, Japanese Journal of Forestry, 15:25-35

Woollons, R.C. and Wood, G.R., 1992, Utility and performance of five sigmoid yield-age functions, fitted to stand growth data, In: Wood, G.B. and Turner, B.J. (editors) IUFRO Conference on 'Integration forest information over space and time', Jan. 1992, Canberra, Australia, 71-79

Xu, L., Wood, G.R., Woollons, R.C. and Whyte, A.G.D., 1992, Stand table prediction with reverse Weibull and extreme value density functions: some theoretical considerations, Forest Ecology and Management, 48:175-178

Yang, R.C., Kozak, A. and Smith, J.H.G., 1978, The potential of Weibull-type functions as flexible growth curves, Canadian Journal of Forest Research, 8(4): 424-431

Zhang, S., Burkhart, H.E. and Amateis, R.L., 1996a, Modelling individual tree growth for juvenile loblolly pine plantations, Forest Ecology and Management, 89(1-3): 157-172

Zhang, S., Burkhart, H.E. and Amateis, R.L., 1996b, A process-oriented individual tree increment model for juvenile loblolly pine plantations, In: Skovsgaard, J.P. and Johannsen, V.K. (editors) The IUFRO Conference on 'Modelling Regeneration Success and Early Growth of Forest Stands', June, 1996, Copenhagen, Denmark, Ministry of environment and energy, Danish forest and landscape research institute, p. 298-305

Zhang, S.A., Amateis, R.L. and Burkhart, H.E., 1997, Constraining individual tree diameter increment and survival models for loblolly pine plantations, Forest Science, 43(3): 414-423

Zhao, W., Guo, L. and Guo, J., 1991, Preliminary study on a new stand density measure, Journal of Shanxi Agricultural University, Forestry Collection, 42-45

Zhao, W. and Mason, E.G., 1996, Early growth modelling of *Pinus radiata* in Canterbury, New Zealand, In: Skovsgaard, J.P. and Johannsen, V.K. (editors) The IUFRO Conference on 'Modelling Regeneration Success and Early Growth of Forest Stands', June, 1996, Copenhagen, Denmark, Ministry of environment and energy, Danish forest and landscape research institute, p. 306-315

APPENDICES

A computer simulation program and two published papers that were prepared during the study are appended.

The computer simulation program of CanSPBL is provided on a floppy diskette and it is confidential to W. Zhao, Dr E. G. Mason and Selwyn Plantation Board Limited for two years. Two files are packed in this program. The file CanSPBL.xls is the main program and it can only be run in Microsoft Excel 97. Its features and construction were described in chapter 8. The file Readme.txt is a guide for installation and removing.

The file CanSPBL.xls can be installed in any path in the user's system. If a reminder for macro check for virus is displayed when the file is opened at the first time within Excel, the command button **Enable Macros** should be clicked. A hidden button used to open the system will remain after its first running. Removing the program permanently from the user's system can be done by manually deleting the file CanSPBL.xls and Readme.txt first and then delete toolbar CanSPBL. To delete the toolbar, users can open Excel, click **Tools** on the menu, click **Customize**, click the **Toolbars** tab, select **CanSPBL**, click **Delete**, click **OK**, and then click **Close**.

Paper 1 (re-print):
Zhao, W., & E.G. Mason, 1996, Early growth modelling of Pinus radiata in Canterbury, New Zealand, IN: J.P. Skovsgaard and V.K. Johannsen (Eds), Proceedings of IUFRO conference on "Modelling regeneration success and early growth of forest stands", Danish Forest and Landscape Research Institute, Copenhagen, Denmark

Paper 2:
Mason, E.G., D.B. South, & W. Zhao, 1996, Performance of *Pinus radiata* in relation to seedling grade, weed control, and soil cultivation in the Central North Island of New Zealand, New Zealand Journal of Forestry Science 26 (1/2): 173-183

EARLY GROWTH MODELLING OF *PINUS RADIATA* IN CANTERBURY, NEW ZEALAND

W. Zhao & E.G. Mason

School of Forestry, University of Canterbury
Christchurch, New Zealand

ABSTRACT

This paper describes the creation of early growth models for tree height, diameter at root collar and survival of *Pinus radiata* D. Don in Canterbury, New Zealand. Seven sets of experimental data were used to build the models and they were developed using three procedures: principal component analysis to select environmental variables, linear multivariate regression to test parameter hypothesis, and non-linear regression to recalibrate parameter estimates with the full data set. Annual rainfall was identified as the major factor in determining the growth rate. Weeding and nitrogen fertiliser also significantly entered the models. More data are needed to cover the wide range of the sites in the region.

INTRODUCTION

Early growth modelling for plantations of *Pinus radiata* D. Don concentrates on the state from immediately after planting to just four or five years after planting. The main management objectives during the early growth are to obtain enough trees with few defects and gain a large tree size.

Early plantation growth can be simply expressed as a function of conditions of seedlings and sites at any stage. Seedling quality can be described physically and morphologically and it can be changed by genetic improvement, nursery regimes and stock handling (Menzies, 1988; Genetics and Tree Improvement Research Field, 1987). Site productivity is determined by location, soil, weather and other factors (Hunter, 1984). The micro-environment can be changed by such field management options as land clearing, cultivation, fertilisation, and weed management (Mason, 1992, 1995; Balneaves, 1982; Hunter, Rodgers, Duningham, Prince, & Thorn, 1991).

A good mathematical form for early growth of tree height or diameter is basically the yield form $Y_T=Y_0+\alpha T^{\beta}$ (Belli, 1987; Belli & Ek, 1988; Mason, 1992). Belli (1987) introduced the modifier function to his base form called potential performance function for only two known management variables. Mason (1992), to more efficiently identify the significant factors among huge numbers of environmental and management variables, used analysis of variance for individual plot parameter estimates created with non-linear regression. Similar methods have been used by some other modellers. There are two weaknesses for the methods by the authors' experience: 1) at least three measurements for each plot were needed to obtain the parameter estimates with the non-linear function. It means a plot had to be dropped if only one or two measurements were available in the306 plot; 2) when each parameter was predicted with several significant variables and when every parameter prediction model was put into the original nonlinear model, it did not seem that all the variables were significant. The more variables involved, the more likely this is. In this study the same model was adopted for height and diameter but the parameter form was modified. Its parameter hypothesis was tested with linear multivariate regression and recalibrated with non-linear regression with the full data set. Both linear and nonlinear procedures were conducted by using direct response variable so the two analyses were compatible.

EXPERIMENTAL DATA

Experiments

There were eleven sets of experimental data available which were serial uniformity trials used to examine the response of cultivation, weeding and fertilisation during establishment. The experiments were located in four main regions: Balmoral Forest, Eyrewell Forest, Hanmer Forest and Mcleans Island Forest. Nine of the experiments were established in the 1970's by the New Zealand Forest Research Institute (FRI) and two others were set up in 1983 by Euan Mason and Patrick Milne. Height and root collar diameter (Rcd) measurements were available mostly up to four years old. Survival could be derived from the assessment. Table 1 lists the information for each experiment.

Table 1 Summary information for each individual experiment

Experiment	location	Age at which data available	Treatments and Levels			Initial stock	Altitude	Annual rainfall during trial
			Cultivation	Fertilization	Weed control			
[1]C393	Hanmer	0,2,4	0, 1	N*P (3*4)	0	1670	400	1446
[1]C394	Balmoral	0-4	0	0,1 N+P+B	1	1667	200	666
[1]C396	Eyrwell	0,2,4	0	N*P (3*4)	0	1250	140	933
[1]C404	Balmoral	0,2,4	0,1	N*P (3*4)	0	1667	290	647
[1]C451	M.Lean.I	2,4,6	0	0,1 P+B	1	1764	75	760
[1]C507/1	Balmoral	1-6	0	0,1 N+P	0,1	625	295	655
[1]C507/2	Eyrwell	0-3	0	0,1 N+P	0,1	667	150	887
[1]C515	Balmoral	0-4	0	0,1 N+P+	1	1000	260	629
[1]C516	Balmoral	0-4	0	N*P*K*S 2^4	1	1000	250	629
[2]C573/1	Eyrwell	0-5	0,1	0,1 N+P+	0,1	1250	135	786
[2]C573/2	Balmoral	0-4	0,1	0,1 N+P+	0,1	1250	240	681

1 established by Forestry Research Institute (FRI) 2 established by Euan Mason

An analysis of variance was calculated for each experiment individually. There were four trials designed to test cultivation and only one of them revealed a significant effect on the growth of height and root collar diameter (Rcd). Weeding, however, was significant in all four designed experiments. The fertilisation involved was mainly nitrogen, phosphorous, potassium, and boron. Nitrogen or combinations of nitrogen with other elements significantly increased height and diameter growth in seven out of ten experiments in which N was a factor. Survival often did not change significantly with the treatments. Two trails showed marginally significant survival reduction with fertilisation and two others with weeding due to Velpar poisoning.

Features of the data for modelling and reorganising of the data

Every experiment was carefully designed, established and measured. The weakness was the data might not sample the full range of site conditions in the region. Some treatments in the experiments were designed uncommon, as could be seen in table 1, so they were dropped from the modelling process. Data were dropped were: C451 and C5071 due to lack of measurements immediately after planting; legume treatments in C507/2; refertilisation in C515 and C516; sulphur fertiliser in C516; C573/1 and C573/2 due to their inconsistency with other experiments established by FRI. It was noticed that there was a gap between the two groups of data source. The trees in experiments established in 1983 by Euan Mason had a higher growth rate than those established in the 1970's. The reason might be a genetic difference but could not be clearly identified.

Weed Control was a binary variable. Fertiliser involved four elements: N, P, K and B. They were quantitative variables according to their amounts. Cultivation was not significant in preliminary analyses so the levels of cultivation were not intended to distinguish. Data were organised and matched according to above variables at establishment. Such statistics as mean and others for each plot were appended from age 0 to age 4 for each experiment. Each observation contained the variables: experiment number, T (age), height, diameter, survival, weed control, N fertiliser, P fertiliser, K fertiliser, B fertiliser and so on.

Other Environmental Data

In order to relate growth functions with all potential environmental variables accounting for growth differences between experiments, further information was located:

a) location factors including altitude, latitude, longitude, and distance from sea, were obtained from Department of Land and Survey Information Maps;

b) soil factors including total organic carbon% (Soil-C), total nitrogen% (Soil-N), phosphorus extracted by 1% citric acid (Soil-P), total exchangeable potassium (Soil-K), exchangeable magnesium (Soil-Mg), Cation-exchangeable Capacity (CEC), percentage base saturation (BS), pH, and Carbon Nitrogen Ratio (CNR), were commonly available from Soil Bureau Bulletin 27 (Dept. of Science and Industrial Research, 1968);

c) weather factors including annual rainfall, annual average temperature and the number of annual ground frosts, obtained from meteorological observations in each forest. Averages during the experimental periods were used;

d) seedling conditions including sturdiness (Height Diameter ratio (Menzies, 1988)), batch average root collar diameter and average height were calculated for each plot. All the trees had Growth and Form (GF) ratings less than 15.

FUNCTIONS

A yield functional form was adopted. Difference forms can hide effects behind the T1, Y1 combination (Mason, 1992, 1996). A sigmoid curve can not be expected for the short period of early growth. A part of the sigmoid function curve may be applicable but there is a danger when future yield is extrapolated. Other types of functions like exponential and polynomial functions may logically fit better. In this study the following types of functions were tried. S represented survival and Y were height or Rcd. T was age, Y_0 initial value at age zero, and α, β parameters.

$$S_T = e^{\alpha T^{\beta}} \tag{1}$$

$$Y_T = Y_0 + \alpha T^{\beta} \tag{2}$$

$$Y_T = Y_0 e^{\alpha T^{\beta}} \tag{3}$$

$$Y_T = Y_0 + \alpha T + \beta T^2 \tag{4}$$

Function (4), a quadratic form considered flexible, was intended to try only for comparison with functions (2) and (3) which can be derived from certain theoretical and biological assumptions on early growth (Mason, 1992). Function (1) behaved well for survival modelling and function (2) was a good basic form for height which were displayed in the research by Belli (1987, 1988) and Mason (1992). When the parameters in the model (2) were assumed a linear combination of total environment and management variables, Mason (1992) used the form

$$Y_T = Y_0 + (\alpha_0 + \alpha_1 V_1 + \alpha_2 V_2 + \ldots \alpha_n V_n) T^{(\beta_0 + \beta_1 V_1 + \beta_2 V_2 + \ldots + \beta_n V_n)} \tag{5}$$

where Y_0 was the initial value at age zero, n was the number of total explanatory variables, v_1, v_2... v_n were explanatory variables, α_1, α_2,...α_n, β_1, β_2...β_n parameters in the model. If the first parameter α in the model (2) is put into e^a and the parameter a is assumed to be a linear combination of total variables, then the model (2) becomes model (6) which can be transformed with natural logarithm into a linear model (7). The test of parameter hypothesis then can be done by using linear regression of direct response variable rather than using analysis of variance of the parameters estimates. In following model, a_1, a_2...a_n, b_1, b_2,...b_n were parameters, and ln was the natural logarithm.

$$Y_T = Y_0 + e^{(a_0 + a_1 V_1 + a_2 V_2 + ... a_n V_n)} T^{(b_0 + b_1 V_1 + b_2 V_2 + ... + b_n V_n)} \quad \textbf{(6)}$$

$$\ln(Y_T - Y_0) = a_0 + a_1 V_1 + a_2 V_2 + ... a_n V_n + b_0 \ln(T) + b_1 V_1 \ln(T) + b_2 V_2 \ln(T) + ... + b_n V_n \ln(T) \quad \textbf{(7)}$$

METHODS AND PROCEDURES

Principal Component Analysis (PCA) was applied to reduce the number of environmental variables and to obtain a set of mutually independent explanatory variables (a similar procedure was applied by Hunter (1984)). The major purposes of principal components analyses are to describe a set of variables in terms of the number of 'dimensions' smaller than the number of variables and to indicate what these dimensions represent. There were 15 environmental variables included in the procedure of PCA. All percentage variables were transformed by an arc sine square root calculation and some others by natural logarithms where necessary.

To test the parameters in the transformed linear model (7), stepwise regressions were done by using direct response variables with a partial data set. Multiple linear regression could deal with a number of explanatory variables in an effective and unique solution while non-linear regression could lead to such problems as a singular Jacobian matrix and inefficiency in identifying significant variables if too many variables were involved. Some measurements including all the ones at age zero were dropped from this phase of the analysis and mostly only one measurement per plot was left in using to reduce dependence of errors due to the repeated measurements (West, 1984, 1995).

To recalibrate the parameter estimates in the selected model non-linear procedures were run with model (6) by using the full data set. It was easy to get NLIN (SAS Institute Inc., 1989) to converge after the number of variables were reduced by the linear regression procedure. The NLIN procedure could often lead to a smaller residual sum of squares and a better residual pattern than linear regression. Non-linear procedures were used in fitting functions first. Non-linear procedures were run again by involving the selected environment variables from the above linear regression.

RESULTS

Principal Components Analyses (PCA) for environment variables

Environment variables of soil, weather and location were analysed with PCA. The results of PCA applied to soil variables was that the first component, with a high 93.6% of the total variation, had a closer relationship to Cation Exchange Capacity (CEC) than to all others (the eigenvector was

0.39) and CEC itself was highly correlated to all other soil chemical items (0.80 was the least correlation coefficient). So only CEC was chosen by this procedure.

The first component in the PCA procedure to location and weather variables plus CEC, with 63.5% of the total variation, was more correlated with altitude (0.42). The second component which accounted for 31.3% of the total variation was mainly related to annual rainfall and annual average temperature by equivalent eigenvectors of 0.52 and 0.54 respectively. Temperature had a closely negative correlation to annual rainfall (-0.98) and it ranged only from 10° C to 11.8°C. Soil CEC had a correlation coefficient with annual rainfall by 0.84. Thus altitude and annual rainfall were finally chosen in the process of PCA to represent the environmental factors. The correlation coefficient between altitude and annual rainfall was 0.3.

Fitting of growth functions

Height equation was obtained first. All functions fitted the individual plot very well. The results of non-linear procedures for the whole data set listed in table 2 showed that function (2), (3), and (4) all led to a small mean square of errors for the data set. Function (2), however, had the smallest residual mean squares. No obvious bias was detected.

The results of testing parameters in height model with linear multivariate regressions were that the effects of annual rainfall, weeding and nitrogen fertiliser were significant in model (7) at 1% level. Table 3 shows the regression summary and parameter estimates. When procedure NLIN was run for model (6) with the full data set, keeping the above three variables and setting the parameter estimates as starting values, those parameters were recalibrated and shown on the right three columns of Table 3. The scatter of residual against predicted value is shown in figure 1 and no obvious bias appeared. Residuals were all within ±0.4 m. The figure 3 shows the model behaviour. Rainfall was the major factor affecting growth. Weeding was more effective than nitrogen fertiliser.

Fig.1 Residual vs Predicted H

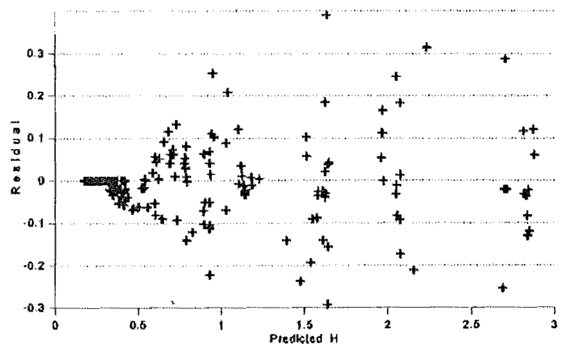
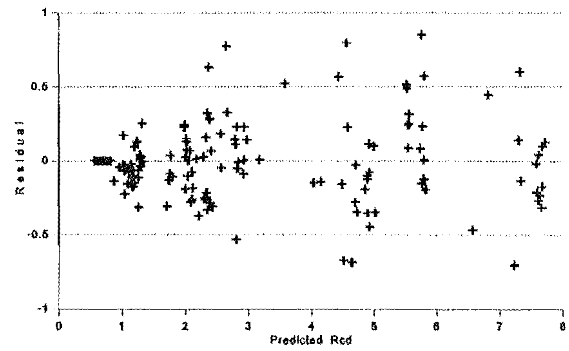


Fig.2 Residual vs Predicted Rcd



The Rcd equation was secondly obtained. The same procedures and the same type of function as above was selected fore Rcd. Significant variables entered into Rcd model were also annual rainfall, weeding and nitrogen fertiliser. Table 4 lists the recalibrated parameters estimates. The model of Rcd behaved similar to that of height. Figure 2 is the scatter of residual against predicted value for the model. No obvious bias appeared. Residuals were all within ± 1.0 cm.

Table 2 Models fitting for three types of functions

Equations	Parameter α Estimate	Std.Error	Parameter β estimate	Std.Error	Residual mean square (RMS)
$Y_T=Y_0+\alpha T^{\beta}$	0.1791872	0.021027	1.6876451	0.089492	0.07376990
$Y_T=Y_0 \exp(\alpha T^{\beta})$	0.5743256	0.041168	0.8935018	0.053330	0.07582347
$Y_T=Y_0+\alpha T+\beta T^2$	0.1023965	0.033872	0.0908905	0.009592	0.07429194

Table 3 Linear & non-linear regression summary and parameter estimates for height

	Dependent Variable: Ln(H) Linear Model (7)			Dependent Variable: H Non-linear Model (6)		
Source	DF	Sum of Squares	Mean Square	DF	Sum of Squares	Mean Square
Model	5	106.66105	21.33221	6	273.00520	45.50
Error	76	1.80987	0.02381	181	1.4939615	0.00825
C Total	81	108.47092	($r^2 = 0.98$)	187	274.49916	
Variables	Para-meter	Parameter Estimate	Std. Error	Para-meter	Parameter Estimate	Std. Error
Intercept	a0	-3.148492	0.071371	a0	-2.833227	0.081296
Annual rainfall	a1	0.000914	0.000056	a1	0.000768	0.000023
Weeding	a2	0.250915	0.049372	a2	0.348215	0.024201
Log(1+Nfertiliser)	a3	0.160778	0.020257	a3	0.142236	0.029248
Log(T)	b0	2.031095	0.054018	b0	1.876607	0.057667
Log(1+Nfertiliser)*log(T)	b1	-0.118928	0.021601	b1	-0.08988	0.022080

Function (1) was used for survival. Its expanded model can be written as function (8)

$$S_T = e^{(a_0+a_1V1+a_2V_2+...+a_nV_n)T^{(b_0+b_1V_1+b_2V_2+...+b_nV_n)}} \qquad \textbf{(8)}$$

where n was the number of total explanatory variables, v_1, v_2... v_n were explanatory variables, a_1, a_2, b_1, b_2,...b_n were parameters in the model.

Parameter estimates were created for each plot with NLIN procedures first. The linear regression of the parameter estimates against independent variables was done and the result was that the second parameter β was correlated to nothing significantly. When parameter β was estimated with full data set as a constant b_0, the model then could be transformed with logarithm into linear model (9). After parameters test for the model with linear multivariate regression of survival against explanatory variables was directly done, annual rainfall and seedling height after planting were indicated the two significant variables. Finally the recalibrates of parameters estimates with non-linear procedure in the full data set were calculated. Table 4 lists the parameter estimates on the right three columns. Figure 4 shows the model behaviour.

$$\ln(S_T) = a_0T^{b_0} + a_1V_1T^{b_0} + a_2V_2T^{b_0} + ... + a_nV_nT^{b_0} \qquad \textbf{(9)}$$

Fig.3 Height Prediction

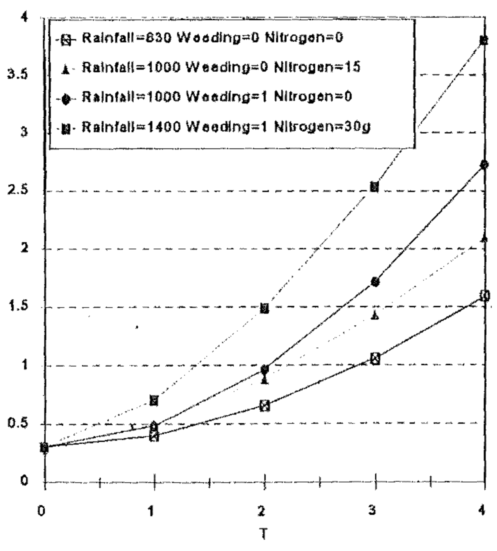


Fig.4 Survival Prediction

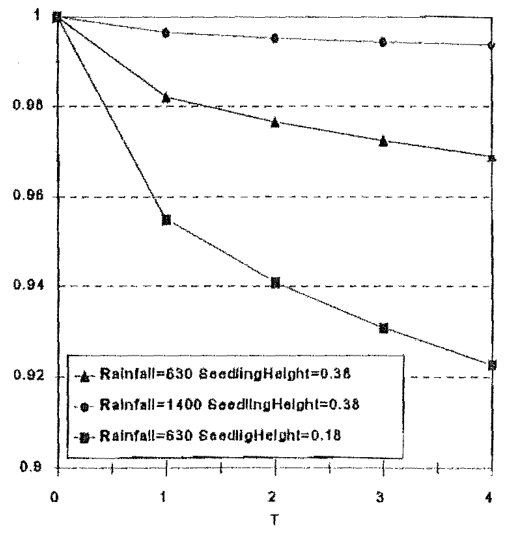


Table 4 Parameter estimates for Rcd & Survival with non-linear regression

Dependent Variable: Rcd Model (6) RMS=0.05536			Dependent Variable: Survival Model (8) RMS=0.00058500		
Variable	Para-meter	Parameter Estimate	Variable	Para-meter	Parameter Estimate
Intercept	a0	-1.795443739	Intercept	a0	-.0831191933
Annual rainfall	a1	0.000652783	Annual rainfall	a1	0.0000187533
Weeding	a2	0.407405713	Seedling height	a2	0.1403957932
Log(1+Nfertiliser)	a3	0.159668493	Intercept	b0	0.401545
Intercept	b0	1.978026412			
Log(1+Nfertiliser)	b1	-0.103219785			

DISCUSSION

The three modelling procedures of principal component analysis to select environmental variables, linear multivariate regression to test parameter hypothesis and non-linear regression to recalibrate parameter estimates with the full data set were very effective for the early growth modelling with a data set of a few measurements for each plot. A function which could not be transformed into linear manner, however, might not be modelled straight forward with the methods, as was done for survival model in this study. The authors agree that the models can be improved for practical use by improving data quantity and quality. Data covering a wide range of sites and data containing various known genetic improvements with consistent treatments would be preferred. Thus we will be able to estimate the gain from genetic improvement. But in this study those two experiments C573/1 and C573/2 had to be dropped due to their unclear genetic class.

CONCLUSIONS

This study showed that functions (1) and (2) were good basic forms for early growth modelling. After the environmental variables and management options were introduced through three procedures of analyses, the models residuals sum of squares were reduced to a minimum. Residuals for height were all within ±0.4 m, for Rcd within ±1.0 cm and for survival within ±10%.

Annual rainfall described growth differences very well among experiments and it was the major factor limiting the growth in the region. Such management options as nitrogen fertiliser and weeding significantly increased growth.

ACKNOWLEDGEMENTS

Acknowledgements are expressed to the former New Zealand Forest Service and the NZ Forest Research Institute for providing data and information.

REFERENCES

Balneaves,J.M.,1982, Grass control -Developing a regime for radiata pine establishment on low rainfall sites, New Zealand Forestry Institute What's New in Forestry Research No.113; 4 pp

Belli, K.L., 1987, FPES: A framework for modelling the artificial regeneration system, IN: Ek, A.R., Shifley & T.E. Burk (Eds.), Proceedings of the IUFRO symposium on forest growth modelling and prediction, USDA, Forest Service General Technical Report NC-120: 361-368

Belli, K.L., & A.K. Ek, 1988, Growth and survival modelling for planted conifers in the Great Lakes Region, Forest Science, 34 (2): 458-473

Department of Science and Industrial Research, 1968, General Survey of Soils of South Island, New Zealand, Soil Bureau Bulletin No. 27, 404 pp

Genetics and Tree Improvement Research Field, 1987, Which radiata pine seed should you use?, New Zealand Forest Research Institute What's New In Forest Research No. 157; 4 pp

Hunter, I.R., & Gibson,A.R., 1984, Predicting *pinus radiata* site index variables from environmental variables, New Zealand Journal of Forestry Science 14 (1); 53-64

Hunter, I.R., B.E. Rodgers, A. Duningham, J.M. Prince, & A.J Thorn, 1991, An atlas of radiata pine nutrition in New Zealand, New Zealand Forestry Research Institute Bulletin No. 165; 24 pp

Mason, 1992, Decision-support systems for establishing radiata pine plantations in the central North Island of New Zealand, Ph.D thesis, University of Canterbury, Christchurch, New Zealand; 243 pages

Mason, E.G., South,D.B. & Zhao, W., 1995, Interactions among seedling grade, weed control, and soil cultivation for *pinus radiata* in the central North Island of New Zealand. In: Second International Conference on Forest Vegetation Management, FRI Bulletin No. 192; pp 107-109

Mason, E.G., South,D.B.& B.R. Zutter, 1996, Classical growth analysis for a computer age, IN: J.P. Skovsgaard and V.K. Johannsen (Eds), Proceedings of IUFRO conference on "Modelling regeneration success and early growth of forest stands", Danish Forest and Landscape Research Institute, Copenhagen, Denmark

Menzies, M.I., 1988, Seedling quality and seedling specifications of radiata pine, What's New in Forest Research No. 171, Forest Research Institute, Private bag, Rotorua, New Zealand

West, P.W., Ratkosky, R.A., and A.W. Davis, 1984, Problems of hypothesis testing of regressions with multiple measurements from individual sampling units, Forest Ecology and Management, 7: 207-224

West, P.W., 1995, Application of regression analysis to inventory data with measurements on successive occasions, Forest Ecology and Management, 71: 227-234

PERFORMANCE OF *PINUS RADIATA* IN RELATION TO SEEDLING GRADE, WEED CONTROL, AND SOIL CULTIVATION IN THE CENTRAL NORTH ISLAND OF NEW ZEALAND

EUAN G. MASON

School of Forestry, University of Canterbury,
Private Bag 4800, Christchurch, New Zealand

DAVID B. SOUTH

School of Forestry, Auburn University,
Auburn, AL 36849-5418, United States

and ZHAO WEIZHONG

School of Forestry, University of Canterbury,
Private Bag 4800, Christchurch, New Zealand

(Received for publication 24 April 1995; revision 30 October 1996)

ABSTRACT

Two experiments were established in the central North Island of New Zealand to examine survival and growth of *Pinus radiata* D.Don in response to weed control and methods of soil cultivation. Fifth- and sixth-year tree height, diameter and survival were examined in relation to (a) initial tree size expressed in various ways, (b) intensity of weed control, and (c) method of soil cultivation. Of four measures of initial seedling size tested, seedling ground line diameter (GLD) was best correlated with tree performance at one site while initial GLD squared × height was most significant at the other. Control of weeds improved tree growth at both sites, and markedly improved survival of trees at the higher altitude site. Analysis of residuals of an initial growth model constructed with data from 27 experiments suggested that stocks of 1/0 seedlings with mean initial root collar diameters of less than 5 mm performed poorly compared with larger 1/0 stocks.

Keywords: seedling quality; competition; soil tillage; fertiliser; establishment.

INTRODUCTION

Performance of young trees in plantations is a function of both site quality and seedling quality (Mason 1992). The effects of site quality have been extensively studied; as modelled by Mason (1992) for *P. radiata* in the central North Island of New Zealand. Seedling quality has also been studied from genetic (Forest Research Institute 1987), physiological (Rook 1971), and morphological (Chavasse 1980) perspectives. Seedling handling during lifting and transport also profoundly influences the performance of seedlings after planting (Trewin & Hunter 1986). Measures of seedling physiology, such as water potential (Balneaves *et al.* 1992), or of actual seedling performance, such as root regeneration potential (Jenkinson *et al.* 1993), have been shown to be reliable predictors of seedling survival and growth. However, root regeneration potential is of little use for selecting quality seedlings (or predicting post-planting performance of the seedlings) when seedlings are planted immediately after lifting, and water potential can be difficult to measure. Consequently, managers in New Zealand commonly assess seedling quality from morphological features such as sturdiness (the ratio of height to diameter at the root collar), root collar diameter (RCD), height, or root/shoot balance (Chavasse 1980; Menzies 1988).

A few studies have included both site and seedling quality as factors. Trewin & Hunter (1986) found the effects of seedling handling and site amelioration were additive in one study. Results of two studies (Balneaves 1989; Albert *et al.* 1980) implied that planting large seedlings could improve growth to a greater extent than controlling weeds. This implication warranted further testing, and was the first subject of the studies reported here.

Ground line diameter (GLD) has been shown to be a good index of seedling quality within any given batch of seedlings, as has the ratio between height and GLD (Chavasse 1980), but these morphological measurements have not been tested as indices of seedling quality to compare batches of seedlings. It is not known whether reduced performance of small seedlings resulted directly from small seedling size or whether small sizes reflected excessive competition in a nursery bed, uneven conditioning practices, or inferior microsites in the nursery which through their impact on other important factors, such as nutrient reserves, caused the smaller trees in a batch to perform badly after planting. A whole batch of seedlings, however, can be reduced in mean GLD simply by being sown later than another batch, and this might not reduce seedling performance. It was not clear, therefore, that results obtained from comparisons between seedlings of different GLD grades within a batch could be applied to comparisons between batches. This was the second question explored during the studies described below.

METHOD

Studies within Seedling Batches

Two studies were located in the central North Island of New Zealand to examine seedling response to weed control, various methods of soil cultivation, and, on one site, to fertiliser treatment with 80 g diammonium phosphate (DAP)/tree. Each study consisted of four blocks.

Experiment R1835/2 was established at Ohakuri at an altitude of 450 m a.s.l. Weeds on the site comprised mainly bracken (*Pteridium esculentum* (Forst. f.) Cockayne) and Cape broom (*Teline monspessulana* (L.) Koch), although there were also patches of grass. Thirteen cultivation treatments and strip spraying with 8 kg hexazinone/ha were examined with a split-plot design.

Experiment R1846 was located at Wainui, at an altitude of 760 m a.s.l. The site was occupied by intensively grazed pasture. One-half of each block was completely sprayed with 5 kg hexazinone/ha. Nine cultivation treatments were randomly located as at Ohakuri, in a cross-over design with weed control, and then the cultivation×weed control plots were divided into subplots for comparison of growth with and without added fertiliser (80 g DAP/tree). Weed control was expected to ameliorate frost, and therefore large contiguous areas free of weeds were required. It was not feasible to cultivate very small lines with a large bulldozer and this, coupled with the requirement for large areas of weed control, was the reason why a cross-over design was employed. The microsite occupied by each tree was classified as a gully, the slope of a gully, or the top area above gullies, because seedlings in the small (1-m-deep) gullies were expected to suffer more frost damage than those above the gullies. These gullies were a natural feature of the site, and had apparently resulted from erosion. There was no evidence, however, that topsoil was any shallower in the gullies, suggesting that they had formed long ago. There was no clear correlation between presence of gullies and treatments, and several gullies crossed the experiment.

In both experiments the cultivation treatments tested included various combinations of ripping to 60 cm, inverted discing, and v-blading (Table 1). Each cultivated plot comprised three 40-m lines. At Ohakuri these plots were split, with weeds controlled in one half and uncontrolled in the other half. At Wainui, the cross-over subplots were split, and 80 g DAP were placed into a slit 10 cm from each tree in one half of each subplot. Trees were spaced at 2 m within lines, and lines were 3 m apart at both sites. Each cultivation plot therefore comprised 60 trees, of which 20 were assessed. At both sites 10 of the measured trees in each plot had weeds controlled, and at Wainui, 10 in each plot received diammonium phosphate.

Sites were planted using 1/0 *P. radiata* seedlings grown from open-pollinated seed orchard seeds. Measurements of GLD and height of trees in each centre row were collected annually at each site, until such time as diameter at breast height outside bark (dbhob) could be measured, after which it was substituted for GLD. Initially 10 trees per subplot were measured at Ohakuri, and five trees per subplot at Wainui. By age 5 years, a total of 906 trees had been remeasured at Ohakuri, and 476 trees were remeasured at Wainui by age 6. Reductions in trees measured were brought about by mortality within the plots.

TABLE 1–Factors and treatments used in the experiments at Ohakuri and Wainui

Factor	Treatment	Ohakuri	Wainui
Weed control	Hexazinone (1 month after planting), slashing	8 kg/ha	5 kg/ha
	Control (no weed control)	Yes	Yes
Cultivation	Control (no cultivation)	Yes	Yes
	Rip	Yes	Yes
	Rip/flat roller	Yes	Yes
	Rip/two inverted discs	Yes	
	Rip/four inverted discs	Yes	Yes
	Rip/six inverted discs	Yes	Yes
	Six inverted discs only		Yes
	Rip/two inverted dises/hourglass-shaped roller	Yes	
	Rip/four inverted discs/hourglass-shaped roller	Yes	Yes
	Rip/six inverted discs/hourglass-shaped roller	Yes	Yes
	Single-pass v-bladed mound	Yes	
	Double-pass v-bladed mound	Yes	
	Double-pass v-bladed mound/rip	Yes	
	Hand cultivation (20×20 cm, 30 cm depth)	Yes	Yes
Fertiliser	80 g DAP/tree, age 0		Yes
	Control (no fertiliser)	Yes	Yes

After 3 years, as weeds began to reinvade the plots treated with hexazinone, spot-spraying with hexazinone and manual slashing were used to further control weed competition.

Analyses of variance of height increment to and dbhob at ages 5 (Ohakuri) and 6 (Wainui) were conducted using SAS software. Subsequently, four (Ohakuri) and three (Wainui) classes of initial GLD, height, GLD squared × height (d^2h), and height:GLD were established, and models were refitted incorporating these seedling classes and their interactions with the site preparation and weed control treatments.

The effects of site amelioration and initial seedling size class on tree survival were analysed using procedure FREQ in SAS software, using chi square as a measure of the significance of treatments and classifications.

Studies between Seedling Batches

Mason (1992) modelled the initial (ages 0 to 5 years) growth of seedlings planted in 27 experiments in the central North Island of New Zealand. The model was sensitive to altitude, weed control, ripping, mounding, fertiliser treatment with 80 g DAP/tree, and initial stocking. The equation used to model height was:

$$H_T = H_0 + \alpha\, T^{\beta}$$

where T = time since planting, H_T = mean height at age T, and α and β were linear functions of altitude (m), weed control (1 = weed control, 0 = no weed control), ripping (1 = ripping, 0 = no ripping), mounding (1 = mounding, 0 = no mounding), and fertiliser (1 = 80 g DAP/ tree after planting, 0 = no DAP).

As the model did not explicitly include seedling size as an independent variable, the effect of initial seedling size on growth could be examined across sites by plotting deviations from the model against mean initial GLDs calculated from the dataset which had been used for estimating the model's coefficients. GLDs of 15 of the 27 batches of seedlings planted in experiments used for constructing the model (Mason 1992) had been measured after planting. The residuals of the height model were plotted against the mean initial GLD of each batch.

RESULTS

Within Seedling Batches

Controlling weeds increased height growth from 5.17 m to 6.13 m (Ohakuri, age 5, $p<0.0007$) and from 4.16 m to 4.74 m (Wainui, age 6, $p<0.0001$). Dbhob was increased from 83 mm to 104 mm (Ohakuri, $p<0.0001$) and from 79 mm to 90 mm (Wainui, $p<0.0003$). The analyses of variance are shown in Tables 2, 3, 4, and 5. Weed control profoundly affected survival at Wainui, especially in gullies ($p<0.001$) where cold air would have settled (Fig. 1), but did not significantly affect survival at Ohakuri.

Cultivation increased survival from 84% to 96% at Ohakuri ($p<0.001$) but did not affect growth significantly when tested against the cultivation×weed control mean square. It was noted that cultivation partially controlled brushweeds, and the effects of weed competition on dbhob were more marked in those treatments which lacked cultivation ($p<0.05$). Cultivation treatments were confounded with a site gradient at Wainui and so the observed

TABLE 2–Analysis of variance of height increment between planting and age 5 at Ohakuri

Source	d.f.	Sum of squares	Mean square	F	P>F
Block	3	880 412.60	293 470.86	4.26	0.0066
Cultivation	12	1 140 703.32	95 058.61	1.38	0.2236
Cultivation×block	33	2 270 795.32	68 811.98		
Weed control (WC)	1	878 470.54	878 470.54	20.83	0.0007
Cultivation×weed control	12	506 099.05	42 174.92	1.79	0.0889
Cultivation×weed control×block	36	850 483.68	23 624.55		
Ground-line diameter class (GLD)	3	541 428.94	180 476.32	13.34	0.0001
Cultivation×GLD	36	534 654.73	14 851.52	1.10	0.3416
Weed control×GLD	3	114 814.58	38 271.52	2.83	0.0411
Cultivation×WC×GLD	36	406 347.76	11 287.43	0.83	0.7366
Cultivation×block×WC×GLD	126	1 691 864.58	13 534.92		
Residual	605	7 169 043.65	11 849.66		
Total	905	20 131 245.63			

TABLE 3–Analysis of variance of diameter at breast height at age 5 at Ohakuri

Source	d.f.	Sum of squares	Mean square	F	P>F
Block	3	33 225.06	11 075.02	4.50	0.0094
Cultivation	12	39 241.65	3 270.14	1.33	0.2494
Cultivation×block	33	81 193.70	2 460.42		
Weed control (WC)	1	47 183.31	47 183.31	50.67	0.0001
Cultivation×weed control	12	22 854.92	1 904.58	2.05	0.0485
Cultivation×weed control×block	36	33 522.08	931.17		
Ground-line diameter class (GLD)	3	27 609.68	9 203.23	13.45	0.0001
Cultivation×GLD	36	29 438.81	817.74	1.19	0.2393
Weed control×GLD	3	2 248.16	749.39	1.09	0.3559
Cultivation×WC×GLD	36	19 250.64	534.74	0.78	0.8037
Cultivation×block×WC×GLD	126	85 550.14	684.40		
Residual	604	341 009.81	564.58		
Total	904	892 260.81			

TABLE 4–Analysis of variance of height increment between planting and age 6 at Ohakuri

Source	d.f.	Sum of squares	Mean square	F	P>F
Weed control (WC)	1	215 625.003 62	215 625.003 62	21.74	0.0001
Block	3	72 067.963 38	24 022.654 46	2.42	0.0756
Cultivation (CULT)	8	407 370.085 06	50 921.260 63	5.14	0.0001
Block×WC×CULT	55	545 400.409 05	9 916.371 07		
D^2H size (DHSIZE)	2	76 418.485 54	38 209.242 77	5.24	0.0072
WC×DHSIZE	2	7 201.581 17	3 600.790 59	0.49	0.6122
Block×WC×DHSIZE×CULT	84	612 869.025 32	7 296.059 83		
Residual	319	2 093 077.598 8	6 561.371 8		
Total	474	4 515 030.631 6			

effects of cultivation are not reported here. This gradient was due to a gentle slope which resulted in greater frost damage on one side of the experiment. The effects of weed control and fertiliser were not confounded with the gradient, and could therefore be tested. Application of DAP influenced neither growth nor survival.

TABLE 5–Analysis of variance of diameter at breast height at age 6 at Ohakuri

Source	d.f.	Sum of squares	Mean square	F	P>F
Weed control (WC)	1	8 035.489 521	8 035.489 521	14.86	0.0003
Block	3	5 638.565 973	1 879.521 991	3.48	0.0219
Cultivation (CULT)	8	11 898.092 708	1 487.261 588	2.75	0.0215
Block×WC×CULT	55	29 735.838 638	540.651 612		
D^2H size (DHSIZE)	2	4 561.307 378	2 280.653 689	5.25	0.0071
WC×DHSIZE	2	609.658 202	304.829 101	0.70	0.4985
Block×WC×DHSIZE×CULT	83	36 034.672 427	434.152 680		
Residual	319	139 824.903 57	438.322 58		
Total	473	252 926.447 26			

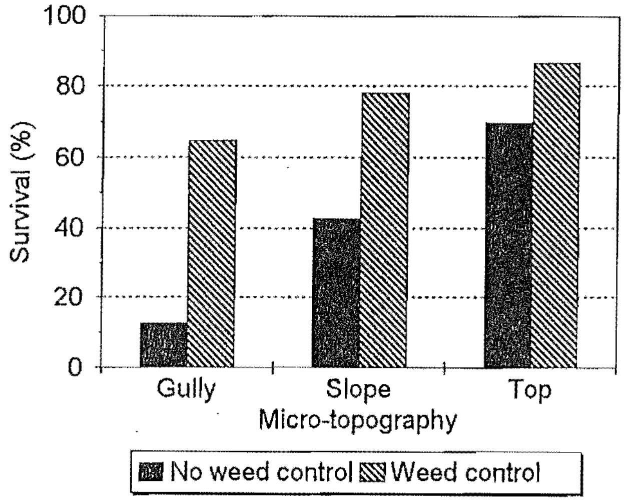


FIG. 1–Survival *v*. weed control and microtopography at Wainui.

GLD was the initial size classification most related to seedling height ($p<0.0001$) and dbh ($p<0.0001$) growth at Ohakuri (Fig. 2). The weed control × GLD interaction was significant for height increment ($p<0.04$). There was a greater increase in height growth with seedling GLD class when weeds were not controlled than when they were controlled. Survival was not significantly related to initial seedling size on this site.

Initial d^2h class was a better indicator of seedling survival ($p<0.001$), height growth ($p<0.0071$), and dbh growth ($p<0.0072$) at the higher altitude site than were the other size classifications (Table 6). Survival was strongly related to microtopography and seedling size (Fig. 3).

Between Seedling Batches

Analysis of residuals of the initial growth model showed a trend with initial GLD. In particular, stocks of 1/0 seedlings with mean initial GLDs of less than 5 mm grew more consistently slowly in height than larger 1/0 stocks (Fig. 4).

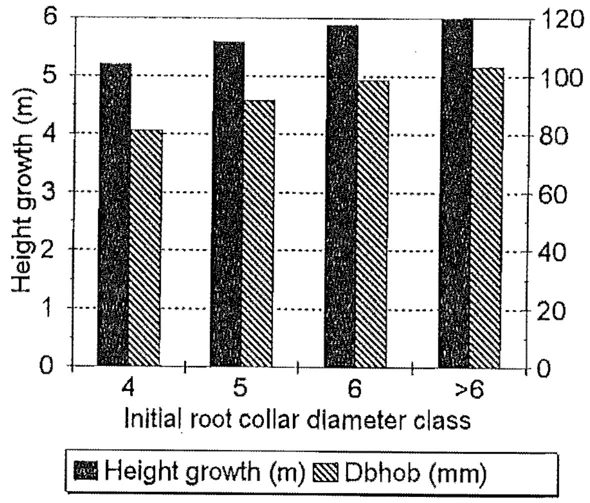


FIG. 2–Height growth and dbhob *v*. GLD at Ohakuri, age 5.

TABLE 6–Effect of initial d^2h class on dbhob and height growth at Wainui (age 6)

Initial d^2h class	Dbhob (cm)	Height growth (m)
400	8.32	4.48
1200	8.74	4.53
2000	8.96	4.61

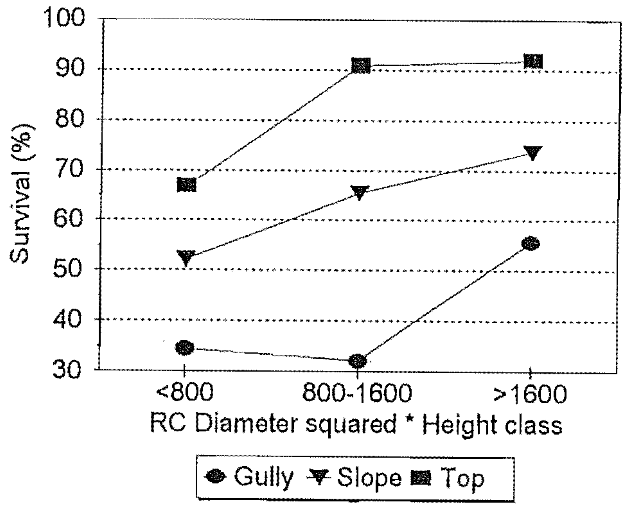


FIG. 3–Survival *v*. ground line diameter squared × height at Wainui.

DISCUSSION

The effects of weed control, cultivation, and fertiliser treatment on growth and survival were consistent with those predicted by the central North Island initial growth model (Mason 1992), and re-emphasise the importance of weed control in this region.

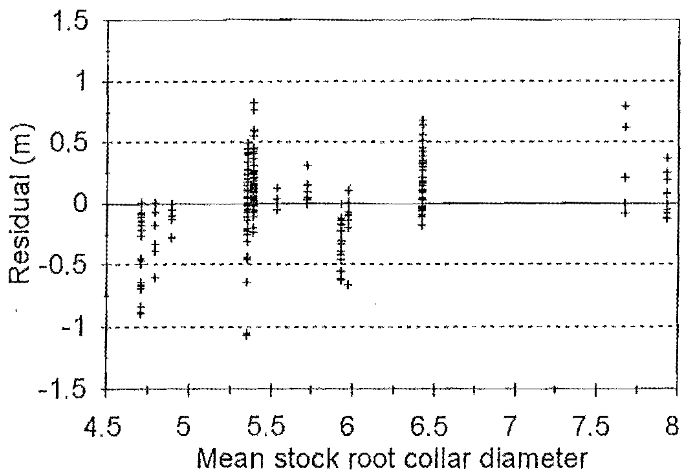


FIG. 4–Residuals of an initial height growth model plotted against the mean initial root collar diameter of seedling batches within experiments used for fitting the model.

The effects of cultivation observed at Ohakuri are consistent with a trend among other experiments which suggests that cultivating light, well-drained soils is unlikely to greatly affect pine height and diameter growth (Somerville 1979; Mason & Cullen 1986; Mason 1992), while cultivating wet, poorly drained ones often results in large growth increases (Hetherington & Balneaves 1973; Williamson 1985; Hunter & Skinner 1986; Mason *et al.* 1993; Derr & Mann 1970; Wilhite & Jones 1981; Haywood 1983; Outcalt 1984). Cultivation with disks was found advantageous, however, on frost-prone central North Island pumice sites where weeds were not controlled chemically, because it incorporated weeds in the soil, thereby exposing mineral soil and raising the albedo of the ground surface (Menzies & Chavasse 1982).

The lack of a response to application of DAP at Wainui is consistent with earlier studies in New Zealand. Phosphorus is applied at establishment mainly on weathered and leached clays or podsolised sands in Northland, or on leached alluvial gravels in the Nelson region. At establishment, nitrogen is sometimes applied in the same localities as phosphorus, but it is also regularly applied to landings throughout New Zealand in attempts to rejuvenate them after logging (Ballard 1978; Will 1985).

Analyses of tree performance after planting suggest that, within a seedling batch, plant size can be a good indicator of potential field performance. Size may represent factors such as root mass, root growth potential, nursery treatment, or genetics which can affect survival and growth. However, size is not an infallible predictor of field performance (especially when comparing seedlings from different nurseries and/or different genotypes). For instance, Mason *et al.* (1988) reported that small (4.7 mm GLD, 23 cm height) 1/0 seedlings performed as well as larger (6.9 mm GLD, 39 cm height) 1.5/0 seedlings from a different nursery. Zwoliński *et al.* (1995) found similar differences between seedling grades within a batch, and concluded that planting large seedlings was economically preferable to controlling weeds. It should be noted that the smaller grade of seedlings in Zwolinski *et al.*'s study (14–20 cm in height and 2.8 mm mean RCD) was much smaller than grades commonly planted in New Zealand. However, a similar conclusion might be drawn from the results reported here if the effects of weed control and seedling size were to be compared by tabulating the differences in performance between the lowest grade with weed control and the largest grade without weed control (Table 7). This sort of inference would be appropriate if managers were actually planting seedlots with only the smallest grade of seedling. However, the seedlings planted at Wainui and at Ohakuri were ungraded, and managers would have had to cull the seedlings in order to increase mean seedling GLD. At both sites, culling small seedlings would have been much less effective than controlling weeds. Effects of culling seedlings less than 5 mm and less than 6 mm on dbhob distributions were compared with effects of weed control using data from Ohakuri (Fig. 5).

TABLE 7–Comparison of effects of weed control with effects of planting only the smallest or the largest seedling grade

Experiment	Initial GLD	Weed control	Height increment (m)	Dbh (cm)	Survival (%)
Wainui	7.5	No	4.5	8.7	66
	3.5	Yes	4.8	8.4	73
Ohakuri	7	No	5.5	9.1	97
	4	Yes	5.7	9.1	92

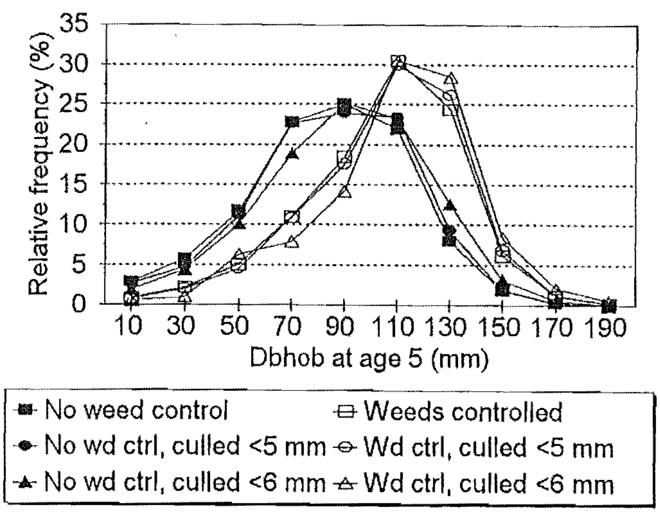


FIG. 5–Consequences of culling seedlings and of weed control on dbhob distributions (age 5) at Ohakuri. The overall mean initial GLD was 5.3 mm. Removing GLD classes less than 5 mm and 6 mm would have resulted in means of 5.6 mm and 6.2 mm, respectively.

The study of residuals from the initial growth model was the first analysis of the effects of GLD measured at planting conducted across a large number of seedling batches in New Zealand. It suggested that managers should aim to plant 1/0 *P. radiata* seedlings with mean GLDs greater than 5 mm. Seedlings planted in the experiments used for model construction were all carefully handled and were planted well (Mason 1992). It is likely that the correlation between GLD and subsequent height growth would have been less evident among batches of seedlings which had been subjected to different levels of abuse during outplanting (Trewin & Hunter 1986).

CONCLUSIONS

Weed control greatly improved growth of *P. radiata* at two sites, and survival at one site where frosts were a problem. Cultivation improved survival at Ohakuri, and facilitated the chemical control of brush weeds. Application of fertiliser at 80 g DAP/tree 1 month after planting made no difference to growth or survival of seedlings planted at Wainui.

Growth and survival were related to initial seedling size expressed as GLD at Ohakuri and d^2h at Wainui. Seedling batches with GLDs less than 5 mm performed poorly compared to other batches in experiments used for the estimation of coefficients during initial growth modelling in the central North Island of New Zealand.

ACKNOWLEDGMENTS

Thanks are due to the former Silvicultural Equipment Development Committee for financing construction of the cultivation rig used during these studies. Permission to use land granted by Carter Holt Harvey Forests (Kinleith) Ltd and Tasman Forestry Ltd was also greatly appreciated. The senior author wishes to thank members of the former Forest Establishment and Equipment research group at the New Zealand Forest Research Institute (NZ FRI) who helped establish the experiment and assisted him with data collection during this study. Thanks also to Noel Davenhill of the NZ FRI for chemical weed control prescriptions and assistance with implementing them.

REFERENCES

ALBERT, D.J.; FRY, G.; POOLE, B.R. 1980: An industrial company's view of nursery stock quality. *New Zealand Journal of Forestry Science 10*: 2–11.

BALLARD, R. 1978: Use of fertilisers at establishment of exotic forest plantations in New Zealand. *New Zealand Journal of Forestry Science 8(1)*: 70–104.

BALNEAVES, J.M. 1989: Root collar diameter of 1/0 radiata pine influences growth following planting. *Forestry Supplement 62*: 125–30.

BALNEAVES, J.M.; MENZIES, M.I.; HONG, S.O. 1992: *Pinus radiata* seedling water potential and root and shoot growth as affected by type and duration of storage. *New Zealand Journal of Forestry Science 22(1)*: 24–31.

CHAVASSE, C.G.R. 1980: Planting stock quality: A review of the factors affecting performance. *New Zealand Journal of Forestry 22(2)*: 144–71.

DERR, H.J.; MANN, W.F. Jr. 1970: Site preparation improves growth of planted pines. *USDA Forest Service Research Note SO-106.*

FOREST RESEARCH INSTITUTE 1987: Which radiata pine seed should you use? *New Zealand Ministry of Forestry, What's New In Forest Research No. 157.*

——1988: Seedling quality and seedling specifications of radiata pine. *New Zealand Ministry of Forestry, What's New in Forest Research No.171.*

HAYWOOD, J.D. 1983: Small topographic differences affect slash pine response to site preparation and fertilisation. *Southern Journal of Applied Forestry 7(3)*: 145–8.

HETHERINGTON, M.W.; BALNEAVES, J.M. 1973: Ripping in tussock country improves radiata growth. *Forest Industries Review 4(12)*: 2–7.

HUNTER, I.R.; SKINNER, M.F. 1986: Establishing radiata pine on the North Auckland podsols. *New Zealand Forestry 31(3)*: 17–23.

JENKINSON, J.L.; NELSON, J.A.; HUDDLESTON, M.E. 1993: Improving planting stock quality—The Humbolt experience. *USDA Forest Service, General Technical Report PSW-GTR-143.*

MASON, E.G. 1992: Decision-support systems for establishing radiata pine plantations in the central North Island of New Zealand. Ph.D. Thesis, University of Canterbury: 301 p.

MASON, E.G.; CULLEN, A.W.J. 1986: Growth of *Pinus radiata* on ripped and unripped Taupo Pumice soil. *New Zealand Journal of Forestry Science 16(1)*: 3–18.

MASON, E.G.; CULLEN, A.W.J.; RIJKSE, W.C. 1988: Growth of two stock types on uncultivated, ripped, and ripped and bedded soil at Karioi Forest. *New Zealand Journal of Forestry Science 18(3)*: 287–96.

MASON, E.G.; MILNE, P.J.; CULLEN, A.W.J. 1993: Establishment regimes for radiata pine on yellow brown earths in Southland. *New Zealand Forestry 37(4)*: 24–9.

MENZIES, M.I.; CHAVASSE, C.G.R. 1982: Establishment trials on frost-prone sites. *New Zealand Journal of Forestry 27(1)*: 33–49.

OUTCAULT, K.W. 1984: Influence of bed height on the growth of slash and loblolly pine in a Leon Fine Sand in Northeast Florida. *Southern Journal of Applied Forestry 8(1)*: 29–31.

ROOK, D.A. 1971: Effect of undercutting and wrenching on growth of *Pinus radiata* D.Don seedlings. *Journal of Applied Ecology 8*: 477–90.

SOMERVILLE, A.R. 1979: Root anchorage and root morphology of *Pinus radiata* on a range of ripping treatments. *New Zealand Journal of Forestry Science 9*: 294–315.

TREWIN, A.R.D.; HUNTER, J.A.C. 1986: A containerised handling system for bare-rooted seedlings. *In* Proceedings of the 18th IUFRO World Congress, Ljubjana, Yugoslavia.

WILHITE, L.P.; JONES, E.P. Jr. 1981: Bedding effects in maturing slash pine stands. *Southern Journal of Applied Forestry 5(1)*: 24–7.

WILLIAMSON, M.J. 1985: Cultivation in Northland. *New Zealand Journal of Forestry 30(2)*: 218–31.

WILL, G.M. 1985: Nutrient deficiencies and fertiliser use in New Zealand exotic forests. *New Zealand Forest Service, Forest Research Institute Bulletin No.97.*

ZWOLINSKI, J.B.; SOUTH, D.B.; CUNNINGHAM, L.; CHRISTIE, S.I. 1995: Weed control and large planting stock improve survival and early growth of *Pinus* radiata. Pp.106–8 *in* Gaskin, R.E.; Zabkiewicz, J.A. (Ed.) "Popular Summaries from Second International Conference on Forest Vegetation Management". *New Zealand Forest Research Institute, FRI Bulletin No.192.*